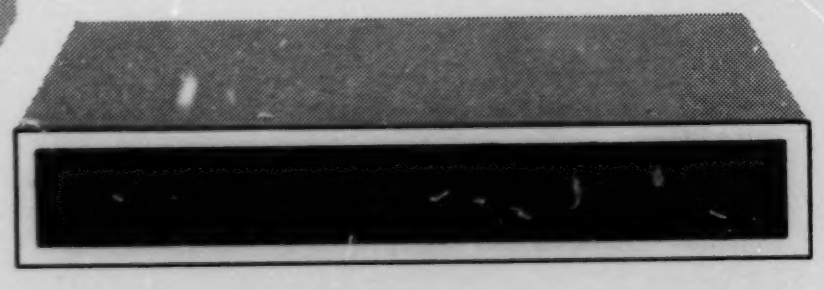
PB 201 408

SAROAD

USERS MANUAL



Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

U. S. ENVIRONMENTAL PROTECTION AGENCY

64

USERS MANUAL: SAROAD
(STORAGE AND RETRIEVAL OF AEROMETRIC DATA)

ENVIRONMENTAL PROTECTION AGENCY
Office of Air Programs
Research Triangle Park, N. C.
July 1971

The APTD (Air Pollution Technical Data) series of reports is issued by the Office of Air Programs, Environmental Protection Agency, to report technical data of interest to a limited number of readers. Copies of APTD reports are available free of charge to Federal employees, current contractors and grantees, and non-profit organizations - as supplies permit - from the Office of Technical Information and Publications, Environmental Protection Agency, P.O. Box 12055, Research Triangle Park, North Carolina 27711 or from the National Technical Information Service, 5285 Port Royal Road, Springfield, Virginia 22151.

Office of Air Programs Publication No. APTD-0663

STANDARD TITLE PAGE FOR TECHNICAL REPORTS	1. Report No. APTD-0663	2. Gov't Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle: Users Manual: Saroad (Storage and Retrieval of Aerometric Data)			5. Report Date: July 1971
			6. Performing Organization Code
7. Author(s)			8. Performing Organization Rept. No.
9. Performing Organization Name and Address: Environmental Protection Agency, Office of Air Programs, Research Triangle Park, N. C. 27711			10. Project/Task/Work Unit No.
			11. Contract/Grant No.
12. Sponsoring Agency Name and Address: Same as above			13. Type of Report & Period Covered
			14. Sponsoring Agency Code
15. Supplementary Notes			

16. Abstracts

SAROAD(Storage and Retrieval of Aerometric Data) is the data-handling system adopted by the Environmental Protection Agency to promote standard methods of handling air quality and related meteorological data. This manual explains in detail how agencies should report air quality and related meteorological data in SAROAD formats for inclusion in the National Aerometric Data Bank.

17. Key Words and Document Analysis. (a). Descripters

Manuals
Data storage
Data retrieval
Data processing
Data recording
Punched cards
Magnetic tapes
Air pollution
Meteorology

17b. Identifiers/Open-Ended Terms

SAROAD(Storage and Retrieval of Aerometric Data) National Aerometric Data Bank

17c. COSATI Field/Group 13 B, 5B

18. Distribution Statement	19. Security Class (This Report)	21. No. of Pages
Unlimited	UNCLASSIFIED	69
	20. Security Class. (This Page) UNCLASSIFIED	22. Price

FORM NBS-357(1-70)

FOREWORD

SAROAD (Storage and Retrieval of Aerometric Data) is the data-handling system adopted by the Environmental Protection Agency and is recommended for use nationwide.

This manual informs state and local air pollution control agencies of how to include their data in the National Aerometric Data Bank. Some agencies will submit data on SAROAD paper forms, some on punched cards in SAROAD format, and some on magnetic tape in SAROAD format. Sections of this manual are addressed to each of these methods of submitting data, and, to the extent possible, each section is self-contained; therefore, certain instructions common to all users are repeated in each section.

SAROAD consists of a set of standard formats for recording validated measurements for different pollutants and associated meteorological observations with various averaging times and a set of standard codes for identifying such items as pollutants, site locations, and methods of sampling and analysis. At its inception, SAROAD entailed six basic forms: a single form for identifying site location and five data forms. In several years of use, state and local agencies have suggested improvements to the site form and several of the data forms. This manual describes in detail how to use these improved forms. Although we would prefer working with these improved forms, agencies may use any of the earlier SAROAD forms to incorporate data into the bank. In the sections that follow, this manual defines the applicability of SAROAD formats to various levels of state and local data-handling capability, and explains in detail the mechanics of recording data in SAROAD format and submitting them to the National Aerometric Data Bank.

Questions, comments, and requests regarding SAROAD may be sent to:

National Aerometric Data Bank
Environmental Protection Agency
P.O. Box 12055
Research Triangle Park
North Carolina 27711

Forms, codes, and other information can also be obtained from this address.

This document was prepared under the direction of Mr. Gerald Nehls. The Environmental Protection Agency extends appreciation to the following agencies for their contributions in reviewing this manual:

Arlington Health Department (Virginia)

Colorado Department of Health

Connecticut State Health Department

Fairfax County Health Department (Virginia)

Los Angeles County Air Pollution Control District

Maryland State Department of Health and Mental Hygiene

New York State Department of Environmental Conservation

Pennsylvania State Department of Health

Vermont Department of Health

LIST OF FIGURES

Figure		Page
1	SAROAD Site Identification Form	8
2	SAROAD Hourly Data Form	12
3	SAROAD Daily Data Form	14
4	Card Format Used for Data Recorded at 1-, 2-, 3-, 4-, 6-, 8-, and 12-Hour Intervals	16
5	Card Format Used for Data Recorded at 24-Hour or Longer Intervals	17
6	Card Format for 1-Hour Interval Data	19
7	Various-Interval Values (Less than 24-Hour) Punched on Cards	20
8	Card Format	24
A-1	Portion of Typical Flow Diagram	39
C-1	SAROAD Multiple-Parameter Data Form	44
C-2	SAROAD Multiple-Station Data Form	45
D-1	SAROAD Composite Data Form	48
E-1	Example of Data Coding	52
F-1	Organization of Files on Tapes	56

LIST OF TABLES

Table		Page
F-1	Minimum Information Required About Tape/File Characteristics	57

CONTENTS

INTRODUCTION 1

FLOW AND TYPE OF DATA 1

RETRIEVING DATA FROM DATA BANK 1

DATA PROCESSING 3

DATA GENERATION 3

DATA RECORDING 3

DATA VALIDATION 4

DATA STORAGE AND RETRIEVAL. 5

DATA ANALYSIS AND INTERPRETATION. 5

RECORDING DATA IN SAROAD FORMAT 7

SITE IDENTIFICATION 7

RECORDING SITE IDENTIFICATION INFORMATION 7

RECORDING AND SUBMITTING DATA ON SAROAD FORMS 11

SAROAD Hourly Data Form (Intervals Less Than 24 Hours) 11

SAROAD Daily Data Form (Intervals Greater Than 24 Hours) . . . 12

SUBMITTING DATA ON PUNCHED CARDS (SAROAD FORMAT) 15

General Instructions 15

Data Intervals of 1 Hour 15

Intervals Longer Than Hourly, Less Than Daily 18

Daily or Longer Intervals 23

Converting Data from SAROAD Forms to Punched Cards 23

SUBMITTING DATA ON MAGNETIC TAPE (SAROAD FORMAT) 26

CONVERTING EXISTING DATA TO SAROAD FORMAT 27

SUBMITTING PAST DATA NOT IN SAROAD FORMAT 29

CODE TABLES 30

APPENDICES

A. DEVELOPING A MINIMAL COMPUTERIZED DATA PROCESSING SYSTEM. 37

B. SITE DESCRIPTION DEFINITIONS. 41

C. OPTIONAL FORMS 43

D. COMPOSITE DATA FORM 47

E. EXAMPLES OF ASSIGNING START-HOUR AND LOCATING FIRST OBSERVATION 51

F. DETAILED INFORMATION FOR SUBMITTING DATA IN SAROAD FORMAT ON MAGNETIC TAPE 55

G. SUBMITTING PAST DATA IN NONSAROAD FORMAT 59

USERS MANUAL: SAROAD (STORAGE AND RETRIEVAL OF AEROMETRIC DATA)

INTRODUCTION

The Environmental Protection Agency (EPA) is promoting standard methods of handling air quality and related meteorological data to facilitate the exchange of technical information needed for air pollution abatement, control, and research. EPA has adopted a standard coding structure and formats known as Storage and Retrieval of Aerometric Data (SAROAD) to implement this standardization. EPA uses the SAROAD system in operation of the National Aerometric Data Bank, which contains air quality and meteorological data (in SAROAD format) supplied by various monitoring facilities across the United States. This data bank is a key element in the developing National Aerometric Data Information Service (NADIS), which is a systems approach to the collection, storage, and retrieval of local, state, and federal aerometric data to insure that the appropriate control agencies receive comparable, accurate, and current data. This manual explains in detail how agencies should report air quality and related meteorological data in SAROAD formats for inclusion in the National Aerometric Data Bank.

FLOW AND TYPE OF DATA

The preferred method of data flow is from local agency to state agency, then directly to the National Aerometric Data Bank. This flow of data keeps states apprised of the air quality throughout their jurisdictions. All agencies are requested to provide data on all pollutants or meteorological parameters monitored at one site for at least 3 consecutive months unless specifically requested for data of shorter time periods. If data are available only for a 3-month period, they should encompass a season or a quarter. Data from observations made more frequently than every hour should be averaged into hourly values.

RETRIEVING DATA FROM DATA BANK

Certain summaries of the data will be returned to the contributing agencies routinely as the data are entered into the data bank. The additional data services that can be supplied through the National Aerometric Data Bank depend in large measure on the active cooperation and participation of state and local agencies. Submitting data in the SAROAD format facilitates ready incorporation of data into the bank. If data are not received in SAROAD format, extra time and effort will be consumed in processing and loading the data into the bank and thus will be diverted from providing data handling services to meet specific agency requests.

DATA PROCESSING

Experience has shown that agencies should place as much emphasis on the processing of air quality data as on actual sampling. As a minimum, one-third of the resources allocated to accumulating valid air quality data must be devoted to data processing; the remaining two-thirds are applied to sampling and quality control. As demands upon total resources increase, some agencies now operating entirely by manual methods may wish to consider at least a minimal computer-based data processing system. Development of such a system is discussed in Appendix A.

An effective data processing effort involves five phases:

1. Data generation.
2. Data recording.
3. Data validation.
4. Data storage and retrieval.
5. Data analysis and interpretation.

These phases of data processing, as they apply to air quality surveillance, are discussed in the paragraphs that follow.

DATA GENERATION

Generating raw data (i.e., unprocessed and unvalidated observations representing levels of pollutants or related information) is the first step in the process of acquiring air quality and related meteorological data. In the past air pollution control agencies have given this phase the major share of attention. Such considerations as sampling method, averaging time, instrument type, and site location have been extensively reviewed. Identification and characterization of the monitoring site, however, have sometimes been neglected. Yet data cannot be properly interpreted unless the conditions under which they were gathered are known. All too frequently, detailed information on sampling sites, instruments, and methods are unavailable even to the agency taking the sample. These omissions seriously reduce the usefulness of the data. Agencies should take great care to provide a complete and exhaustive description of the sites (and their surroundings) at which data are gathered, and to provide complete and accurate location data along with descriptions of the sampling methods and analytical procedures. This information about the sampling sites will help assure that data gathered at those sites are not misinterpreted or misunderstood.

DATA RECORDING

With certain instrumentation, data representing pollutant concentrations, the time of occurrence, and other related information are recorded at the sampling location. These data may be in machine-readable or non-machine-readable form or both (e.g., pollutant observations on punched papertape with the information describing instrument maintenance hand-recorded). With other devices such as manual samplers, the only data recorded at the site are identification and

PRECEDING PAGE BLANK

maintenance information. The collected samples are chemically analyzed and the observations recorded at a central location.

Regardless of how the data arrive for processing, at this point the question of format arises. Choice of a format need not present difficulties, however, if one realizes that it simply involves selecting a systematic way of recording data for processing. Selection of a format need not constrain an agency in any way, particularly if data are recorded in machine-readable form, for then the data can be reproduced with very little difficulty into a variety of forms and formats. For example, an agency could easily use one format for certain internal records and another format (SAROAD) for interchange of data with other agencies and with EPA.

The availability of data in a clearly defined and well-documented format is the very basis of an effective data processing system. All too often, data are recorded in a format that is complicated, confused, incomplete, and insufficiently documented. When such data leave the immediate care of the person responsible for recording them, their utility is effectively lost. The SAROAD format provides a complete, systematic method for recording data for storage or reporting purposes. Because a complete description of this format is available to all users, data recorded in this format are immediately accessible and exchange of data among agencies is facilitated. Thus, the SAROAD format is strongly recommended for use by state and local agencies, particularly for interchange of data with other organizations.

DATA VALIDATION

An agency's routine use of an effective data evaluation process determines the validity and usefulness of the numbers recorded. The first major step in data validation is applying corrections for known instances of instrument malfunction, drift, or other deviation from normal instrument operation. Corrections applied during this data validation step are based on:

1. Calibration of all instruments by standard methods.
2. Analyses by standard or recommended procedures.
3. Conduct of a quality control program.
4. Validation of sensor output.

The second major step in data validation is investigation of other possibly incorrect data. The data under scrutiny should be carefully compared with a previously established table of limits for:

1. High values.
2. Low values.
3. Extreme changes in values for consecutive observations.

Limits for high values could be based on air quality standards or on past history of the sampling site. Limits for low values could represent the minimum detectable capability of the instruments and methods of analysis. A low-value check could also be made for consecutive readings below a designated value. Where past data from the site are available, the current values can also be compared with earlier trends as a validation check.

DATA STORAGE AND RETRIEVAL

Systematic data storage provides a means of protecting records against physical deterioration and also insures that past data can be quickly and easily retrieved. Effective security measures must be determined to guard against premature release of data before they have been thoroughly validated. Data retrieval involves procedures whereby past data can be readily located, retrieved, used, and returned to their proper place in storage. Careful and complete documentation of the location and content of stored records will insure that established procedures are followed and records are used properly.

DATA ANALYSIS AND INTERPRETATION

Data analysis usually includes calculating arithmetic and geometric means and their respective standard deviations. The means provide a useful measure of central tendency or the "average" value of a pollutant concentration, while the standard deviations and maximum pollutant concentration are useful for characterizing the spread of the data values about the mean value. Many agencies also calculate frequency distributions, which give the percentage of sampling intervals during which any given pollutant concentration is exceeded. Agencies also may perform more elaborate analyses for research studies or other special uses. Upon completion of data analysis, an effort should be made to interpret the results in terms of meteorological conditions, characteristics of emission sources, and topographic conditions. Data interpretation requires complete and detailed knowledge of the sampling site and the surrounding area, as well as of the instruments and sampling methods. Here again is reflected the importance of the site identification data, which can pinpoint the location of unusual occurrences that may call for even more detailed information about the area. In this way agency personnel can develop an understanding of the characteristics of air quality within their jurisdiction. This understanding is then applied in working with local industries and citizens toward effective prevention, control, and abatement of air pollution.

RECORDING DATA IN SAROAD FORMAT

Air quality and meteorological data can be recorded and submitted to the National Aerometric Data Bank in three physical forms: paper SAROAD forms, cards punched in SAROAD format, or magnetic tape written in SAROAD format. Before describing how to record and submit data in each of these forms, we will consider the preparation of site identification information. This information must be provided for the National Aerometric Data Bank before submission of any data.

SITE IDENTIFICATION

Detailed description of the location and surroundings of the sampling site is a prerequisite for useable data. The SAROAD Site Identificatien form is designed to provide this information. A form will be submitted for each sampling site regardless of whether the data from that site are sent on paper forms, punched cards, or magnetic tapes.

Release of data without careful specification of where and how they were gathered may lead to serious misinterpretations regarding local air quality. Steps are being taken to insure that data are not released unless accompanied by an adequate description of the monitoring site and sampling method. Agencies submitting data to the National Aerometric Data Bank can prevent possible misinterpretations (1) by submitting complete site identification data (on SAROAD Site Identification forms) and (2) by submitting updated site identification data whenever sampling locations or their surroundings change.

A new SAROAD Site Identification form (Figure 1) provides for all the basic descriptive information required concerning each sampling site. One form should be completed for each of an agency's sites.

RECORDING SITE IDENTIFICATION INFORMATION

In preparing the Site Identification form, leave all blocks on the right side blank. EPA personnel will use these blocks in processing the form for entry into the data bank. After entering the name of the state in which the site is located, list the purpose of the sampling site under "Project," on the same line. Code Table 2* provides a listing of projects. Usually this entry would be population-oriented sampling. Under "City Name" enter the name of the area in which the site is located, if the population of that area is 2,500 or more. If the population is less than 2,500 (i.e., if the site is outside the jurisdiction of an area with a population of 2,500 or more), enter the name of the county in which the site is located. In the next line fill in "County Name" even

---

*Code Tables 1 through 6 list the symbols used in SAROAD formats. Because these are reference tables cited many times throughout this manual, they are presented together at the end of the text.

PRECEDING PAGE BLANK

ENVIRONMENTAL PROTECTION AGENCY
National Aerometric Data Bank
P.O. Box 12055
Research Triangle Park, N.C. 27711

Form Completed By _____ Date _____ New ☐ Revised ☐

TO BE COMPLETED BY THE REPORTING AGENCY

State _____ Project _____

(14 36) City Name (23 characters) _____

(37 51) County Name (15 characters) _____

City Population (right justified) [52 53 54 55 56 57 58 59]

Longitude Deg. Min. Sec. [0 0] (60 61) [W] (62 63 64 65 66 67 68 69)

Latitude Deg. Min. Sec. [N] (70 71 72 73 74 75 76)

UTM Zone [60 61] Easting Coord., meters [62 63 64 65 66 67 68 69] Northing Coord., meters [70 71 72 73 74 75 76]

(18 78) Supporting Agency (61 characters) _____

Supporting Agency, continued _____

* * * * * * * *

(14 79) Optional: Comments that will help identify the sampling site (132 characters) _____

(14 79) _____

(14 38) Abbreviated Site Address (25 characters) _____

(over)

DO NOT WRITE HERE

A: State Area Site [1 2 3 4 5 6 7 8 9 10] Agency [11] Project [12 13] Region [77] Action [80]

B: State Area Site [1 2 3 4 5 6 7 8 9 10] Agency [11] Project [12 13] SMSA [14 15 16 17] Action [80]

C: State Area Site [1 2 3 4 5 6 7 8 9 10] Agency [11] Project [12 13] Action [80]

D: State Area Site [1 2 3 4 5 6 7 8 9 10] Agency [11] Project [12 13] Action [80]

E: State Area Site [1 2 3 4 5 6 7 8 9 10] Agency [11] Project [12 13] Action [80]

Figure 1. SAROAD site identification form (front).

TO BE COMPLETED BY THE REPORTING AGENCY

DO NOT WRITE HERE

(14 54) Sampling Site Address (41 characters)

Address, continued

State Area Site

F | | | | | | | | | |

1 2 3 4 5 6 7 8 9 10

Agency 11

Project 12 13

Check the ONE major category that best describes the location of the sampling site.

Next, check the subcategory that best describes the dominating influence on the sampler within approximately a 1-mile radius of the sampling site.

1. ☐ CENTER CITY
   - ☐ 1. Industrial
   - ☐ 2. Residential
   - ☐ 3. Commercial
   - ☐ 4. Mobile

2. ☐ SUBURBAN
   - ☐ 1. Industrial
   - ☐ 2. Residential
   - ☐ 3. Commercial
   - ☐ 4. Mobile

3. ☐ RURAL
   - ☐ 1. Near urban
   - ☐ 2. Agricultural
   - ☐ 3. Commercial
   - ☐ 4. Industrial
   - ☐ 5. None of the above

4. ☐ REMOTE

Station Type 55 56

County Code 57 58 59 60

AQCR Number 61 62 63

AQCR Population 64 65 66 67 68 69 70 71

Specify units

Elevation of sampler above ground

Elevation/Gr 72 73 74

Specify units

Elevation of sampler above mean sea level

Elevation/MSL 75 76 77 78

Time Zone 79

Action 80

Circle pertinent time zone: EASTERN CENTRAL MOUNTAIN PACIFIC YUKON ALASKA BERING HAWAII

Figure 1 (continued). SAROAD site identification form (back).

if the county name was entered in the line above. The parenthetical "23 characters" and "15 characters" indicate the maximum number of letters and numbers including spaces that can be used in each entry. Usually there is enough room to enter the names in full. If there is not (i.e., if the name contains more characters than the number in parentheses) enter a reasonable abbreviation that will fit within the allowable number of spaces.

For the next entry, "City Population," enter the population of the town or city in which the site is located, as given in the 1970 Census. If this figure is not available, list the source from which the population figure was taken. Enter the figures so that the last digit in the figure is in the last block (59). If the population is small enough that the first few blocks are empty, enter zeros in those blocks. If the name of the county was used for "City Name," enter all zeros in blocks 52 through 59.

The next set of blocks identifies the geographic location of the site in either Universal Transverse Mercator (UTM) coordinates or in latitude and longitude, but not both. The first line is for recording latitude and longitude to degrees, minutes, and seconds. Longitude is recorded in blocks 62 through 69, with a W pre-printed in column 62. If degrees of longitude are less than 100°, enter a zero in block 63. The second line is for recording data in UTM Coordinates. With this system, the lower left coordinate is used as reference. The two-digit UTM zone is entered in the first two blocks (60, 61). Then the easting and northing coordinates are read in meters from the appropriate section of the map. Enter the digits so that the last one is in the far-right block, and enter zeros in any empty blocks to the left.

Under "Supporting Agency," enter the name of the city, county, regional, state, or other agency that operates the site. This is the name to which any questions will be directed.

The next several lines are provided for comments that might help characterize the site. Such comments might include "site within two blocks of heavy construction, 1970" or "site on building adjacent to main street of town."

EPA will enter the "Abbreviated Site Address" with an abbreviation appropriate for use in computer reports.

"Sampling Site Address" is the first entry on the back of the form. When possible, enter the full street address, with other identifying information as space allows. For remote or rural sites that may not have a street address, enter a descriptive phrase such as the name of a building or area that will adequately locate the site. Do not repeat city (or county) and state names already entered. The next section is designed to provide more specific information on location of the site. Check one of the four blocks of major descriptor categories, then check the box within that category that best describes the dominant influence from surroundings. Complete definitions of categories are given in Appendix B. Questions may be referred to the address given in the foreword.

For entering elevation of the site above ground level and above mean sea level, meters are the preferred unit. You may enter these figures in any convenient unit, however, as long as the unit is specified.

Finally specify the time zone in which the sampling site is located. All data are stored in the standard time of the area. Circle one of the following:

. EASTERN	. YUKON
. CENTRAL	. ALASKA
. MOUNTAIN	. BERING
. PACIFIC	. HAWAII

All data will then be stored in the standard time of that time zone in the National Aerometric Data Bank.

Mail the completed Site Identification forms to:

National Aerometric Data Bank
Environmental Protection Agency
P.O. Box 12055
Research Triangle Park
North Carolina 27711

New forms are needed only when the site environment changes significantly. Periodically EPA will mail site description listings to participants for correcting or updating.

RECORDING AND SUBMITTING DATA ON SAROAD FORMS

In the past, data were submitted to the National Aerometric Data Bank on forms and in formats described in an earlier SAROAD System Description and Data Coding Manual.* In response to many constructive criticisms and suggestions from state and local agencies, new forms have been prepared to provide greater flexibility. Although the earlier forms are still useful and acceptable, use of the new forms is encouraged. The National Aerometric Data Bank will continue to process data submitted on the older forms.

Of the two new SAROAD forms, the Hourly Data Form is designed for sampling intervals of less than 24 hours, and the Daily Data Form for sampling intervals of 24 hours and greater. Two optional forms may be useful to some agencies in recording data for sampling intervals of 24 hours and greater, as explained in Appendix C. For the reporting of "composite" data (see Appendix D), a separate form is available. As the appendix explains, "composite" data result from analyses performed on a number of individual samples that have been chemically combined. With the changing data requirements, these forms may not always be adequate for future use, in which case new or modified forms will be developed. As long as the SAROAD codes are used, the number of SAROAD forms can be expanded. For simplicity, we recommend the use of either the Hourly or Daily Data Form.

SAROAD Hourly Data Form (Intervals Less than 24 Hours)

The Hourly Data Form (Figure 2) is used for recording data observed or averaged at intervals of less than 24 hours. For example, it would be used for 2-hour coefficient of haze (COH) readings, and for 1-hour gaseous pollutant readings. Since there is space for only 12 values for each day, two forms are required for an entire month's data when 1-hour readings are recorded. For any other interval, an entire month's data fit on one form. Entries on the upper left of the form provide identification.

*USDHEW, NAPCA, APTD 68-8, August 1968.

LESS THAN 24-HOUR SAMPLING INTERVAL

1 Agency

City Name

Site Address

Project

ENVIRONMENTAL PROTECTION AGENCY
National Aerometric Data Bank
P. O. Box 12055
Research Triangle Park
North Carolina 27711

Parameter observed | Method

Time interval of obs. | Units of obs.

State (2-3) Area (4-7) Site (8-10)

Agency (11) Project (12-13) Time (14) Year (15-16) Month (17-18)

Parameter code (23-27) Method (28-29) Units (30-31) DP (32)

Day	St	Hr	Rdg 1	Rdg 2	Rdg 3	Rdg 4	Rdg 5	Rdg 6	Rdg 7	Rdg 8	Rdg 9	Rdg 10	Rdg 11	Rdg 12
19-20	21-22	33-34	35-36	37-38	39-40	41-42	43-44	45-46	47-48	49-50	51-52	53-54	55-56	57-58	59-60	61-62	63-64	65-66	67-68	69-70	71-72	73-74	75-76	77-78	79-80

Figure 2. SAROAD Hourly Data Form.

"Project" refers to the type of program under which sampling is conducted. "Parameter Observed" refers to the pollutant sampled or the meteorological parameter measured, such as "Sulfur Dioxide" or "Ambient Temperature." "Method" refers to the instrument type and/or the analytical method. Although "Time Interval of Observations" is largely self-explanatory, (1 hour, 2 hours, etc.) a more detailed definition of "time interval" is given in Code Table 3. Parameter values should be reported only in those units listed in Code Table 5. On the top right-hand side of the form enter the SAROAD code numbers in the appropriate blocks.

EPA will assign codes for the first line of blocks to the reporting agency when Site Identification forms are initially submitted. They consist of a two-digit code for state, a four-digit code for the area of the state in which the sampler is located, and a three-digit number specifically identifying the site.

In the second line of blocks, a code letter is entered to identify the type of agency submitting the data (see Code Table 1). Project codes identifying type of sampling program are listed in Code Table 2. "Time" is a one-character code used for the time interval of observations (Code Table 3). For "Year" enter the last two digits of the year in which the observation was taken. Months are designated 01 (January) through 12 (December). Codes for "Parameter" and "Method" are given in Code Table 4. EPA maintains a detailed list of the less common pollutants and methods not given in the appendix. Two-digit codes representing the units in which the values are expressed are given in Code Table 5. The final code block, "DP," stands for decimal point; this block designates the number of places to the right of the decimal point in the value entries. For example, to enter a reading of 0.025 PPM, enter "3" in the DP block and then enter "0025" in the value column. The one DP code applies to all values entered on the form.

On the body of the form, the two-block first column, "Day," is the calendar day of the month (e.g., 01, 02). "ST HR" (start hour) calls for a two-digit number signifying the hour the sampling began ("00" for midnight through "23" for 11 p.m.). For specific sampling intervals, go to Code Table 6 to determine start hour. The next 12 four-block columns are for the data values. Two-hour observations for a one-month period fill one sheet. Observations taken less frequently do not fill all blocks of the sheet. For example, data covering observations made every 4 hours for one month fill only the first six sets of data blocks across the page. As noted earlier, listing 1-hour observations for a month requires two copies of the Hourly Data Form. For 1-hour readings, use two lines for each day's observations. The first line gives "00" (midnight) for "ST HR" and lists the a.m. observations. The next line gives "12" (noon) for "ST HR" and lists p.m. observations. Data for the first 15-1/2 days of the month are given on the first sheet, and for the remaining days on the second sheet. For missing data, leave the blocks blank or enter "9999". Specific examples of assigning start hours and positioning the first sample observation are given in Appendix E.

SAROAD Daily Data Form (Intervals Greater Than 24 Hours)

The Daily Data Form (Figure 3) is a month's record of up to four different parameters observed at a common site over periods of 24 hours or longer. This form is useful for 24-hour Hi-Vol or gas bubbler data that are observed daily, every third day, every sixth day, or any other random schedule. One form can be used for up to four parameters as long as the site, time interval, and starting hour are the same.

ENVIRONMENTAL PROTECTION AGENCY
National Aerometric Data Bank
P.O. Box 12055
Research Triangle Park, N.C. 27711

24-HOUR OR GREATER SAMPLING INTERVAL

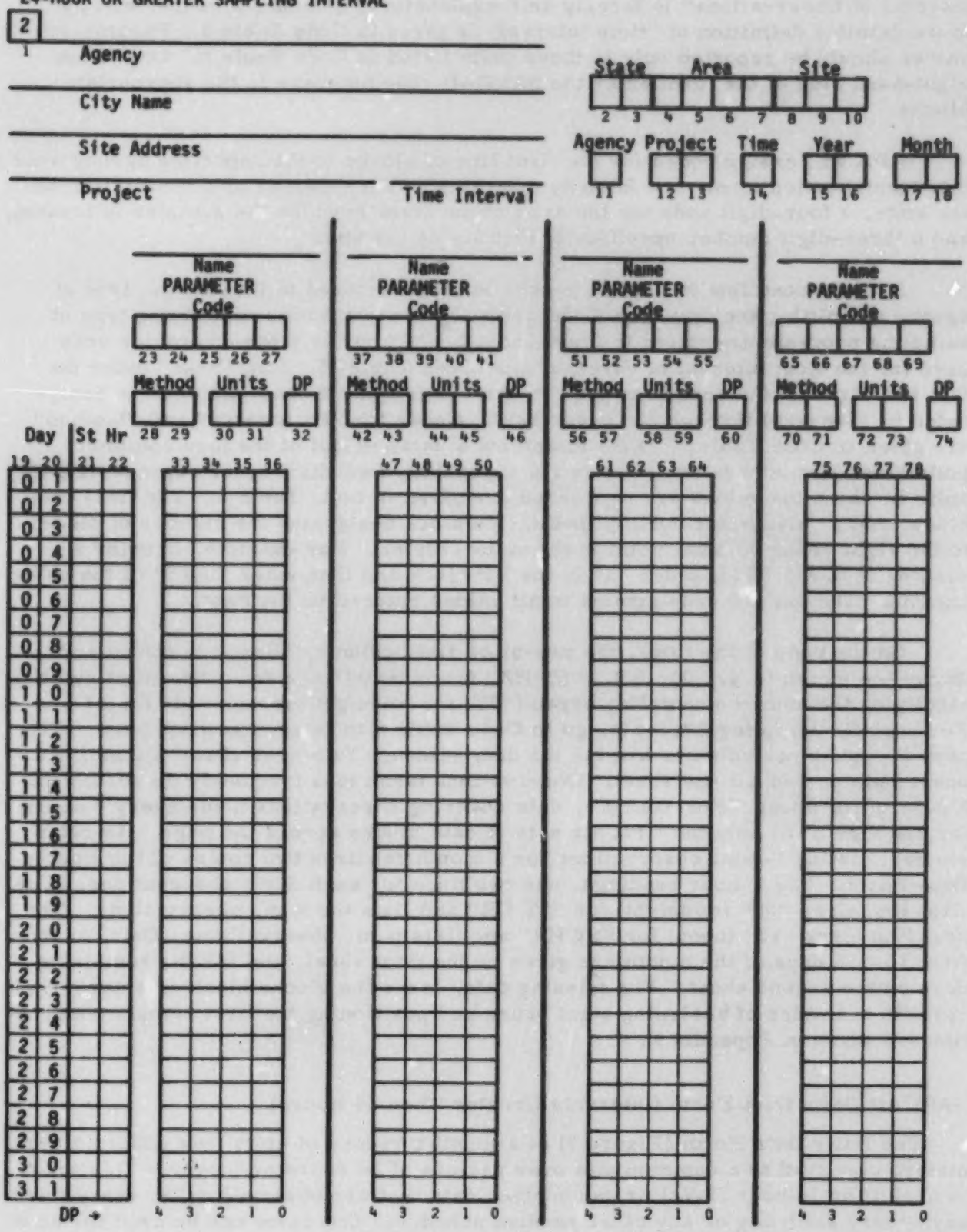


Figure 3. SAROAD Daily Data Form.

Enter the identifying information and codes at the top of Form 2 as on Form 1. Form 2 provides a separate column for each parameter. "ST HR" applies to all data values entered on the same line as the day of the month the observations were made. As before, report missing data by leaving the blocks blank or entering "9999."

Mail copies of the completed forms to the National Aerometric Data Bank at quarterly or more frequent intervals.

SUBMITTING DATA ON PUNCHED CARDS (SAROAD FORMAT)

Many agencies now have the capability for storing data on punched cards, and many other agencies now using paper forms anticipate using automated data processing. Punched cards can be prepared from SAROAD Hourly and Daily Data Forms as explained below.

General Instructions

The first step in recording air quality data on SAROAD-formatted cards is to obtain from the Code Tables the Identification codes for the pollutants monitored, instruments, methods, time intervals, and the like. Data recorded at 1-, 2-, 3-, 4-, 6-, 8-, or 12-hour intervals are punched in the card format shown in Figure 4. Data recorded at intervals of 24 hours or longer are punched in the card format shown in Figure 5. No provision is made for storing data taken at intervals more frequent than hourly. Measurements taken over shorter intervals should be averaged into 1-hour values for reporting.

Certain items of information are punched in the same columns, regardless of the time interval of the observations. The state, station, and site numbers are the same as those that appear on the Site Identification form (Figure 1). EPA assigns site identification numbers to stations reporting for the first time. The identification code "agency type" is punched in column 11. Project classification codes are given in Code Table 2, time-interval codes in Code Table 3. "Year" is indicated by the last two digits of the year in which the observation was taken. Months are punched 01 (January) through 12 (December). Days are designated 01 through 31.

The contents of the remaining columns vary according to the time interval of observations. The following sections specify punching instructions according to the various intervals.

Data Intervals of 1 Hour

Pollutant observations taken at 1-hour intervals or averaged to 1-hour intervals should be punched in the format shown in Figure 4. The format of this card is identified by punching a "1" in the first card column. Columns 2 through 13 are punched according to the general instructions in the preceding section. Punch a "1" in column 14 to identify the time interval of observations as one hour. Since only 12 observations can be punched on a single card, 2 cards are required, one for a.m. and one for p.m. observations. On both cards, columns 1 through 20 will contain the same information. The a.m. and p.m. entries are differentiated by punches in columns 21 and 22, "00" for a.m. and "12" for p.m. Codes to identify parameter and sampling method are given in Code Table 4; codes for units, in Code Table 5. These values will also appear on both cards. Position of the

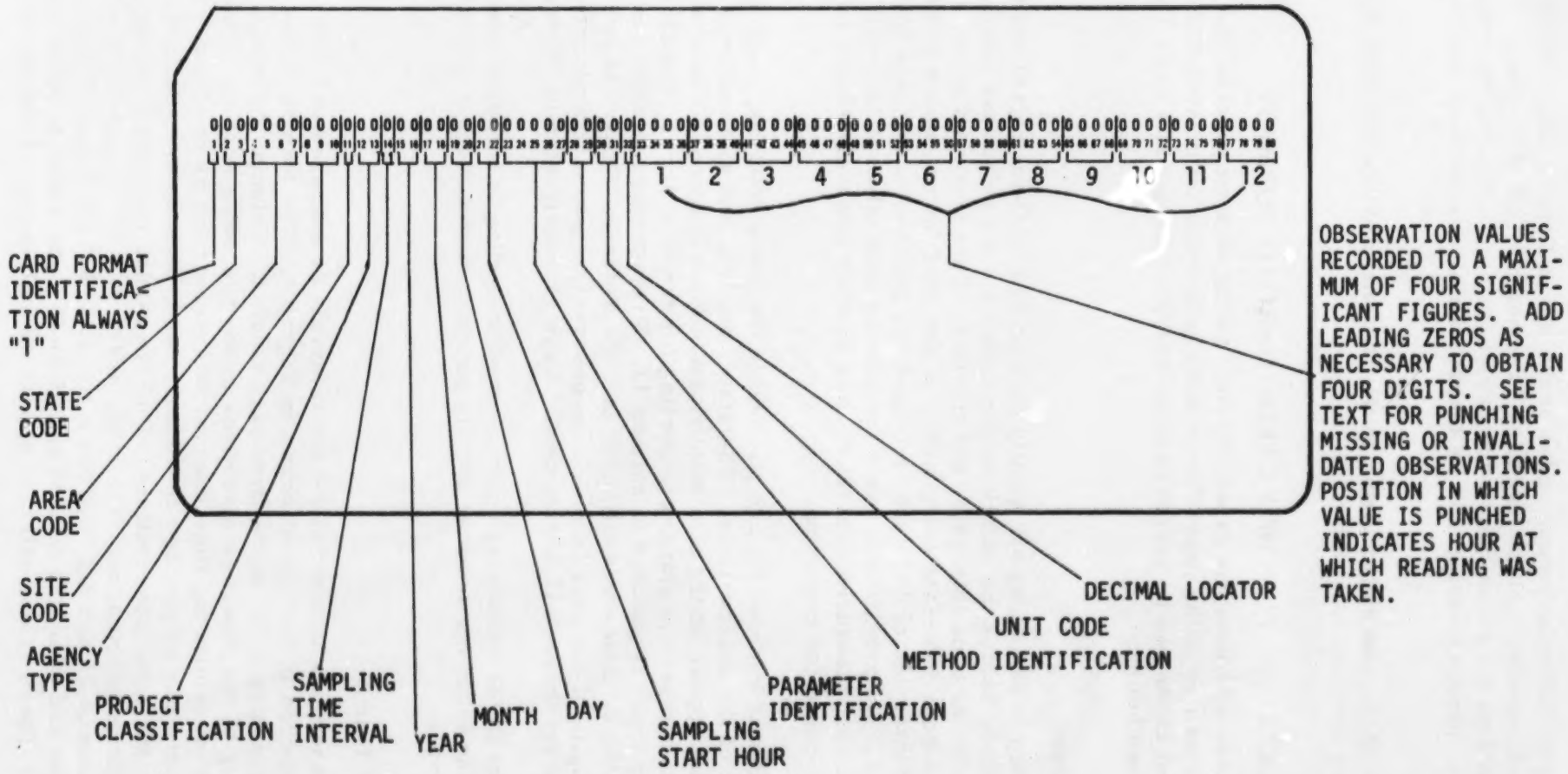

1 2 3 4 5 6 7 8 9 10 11 12
CARD FORMAT IDENTIFICATION ALWAYS "1"
STATE CODE
AREA CODE
SITE CODE
AGENCY TYPE
PROJECT CLASSIFICATION
SAMPLING TIME INTERVAL
YEAR
MONTH
DAY
SAMPLING START HOUR
PARAMETER IDENTIFICATION
METHOD IDENTIFICATION
UNIT CODE
DECIMAL LOCATOR
OBSERVATION VALUES RECORDED TO A MAXIMUM OF FOUR SIGNIFICANT FIGURES. ADD LEADING ZEROS AS NECESSARY TO OBTAIN FOUR DIGITS. SEE TEXT FOR PUNCHING MISSING OR INVALIDATED OBSERVATIONS. POSITION IN WHICH VALUE IS PUNCHED INDICATES HOUR AT WHICH READING WAS TAKEN.

Figure 4. Card format used for data recorded at 1-, 2-, 3-, 4-, 6-, 8-, and 12-hour intervals.

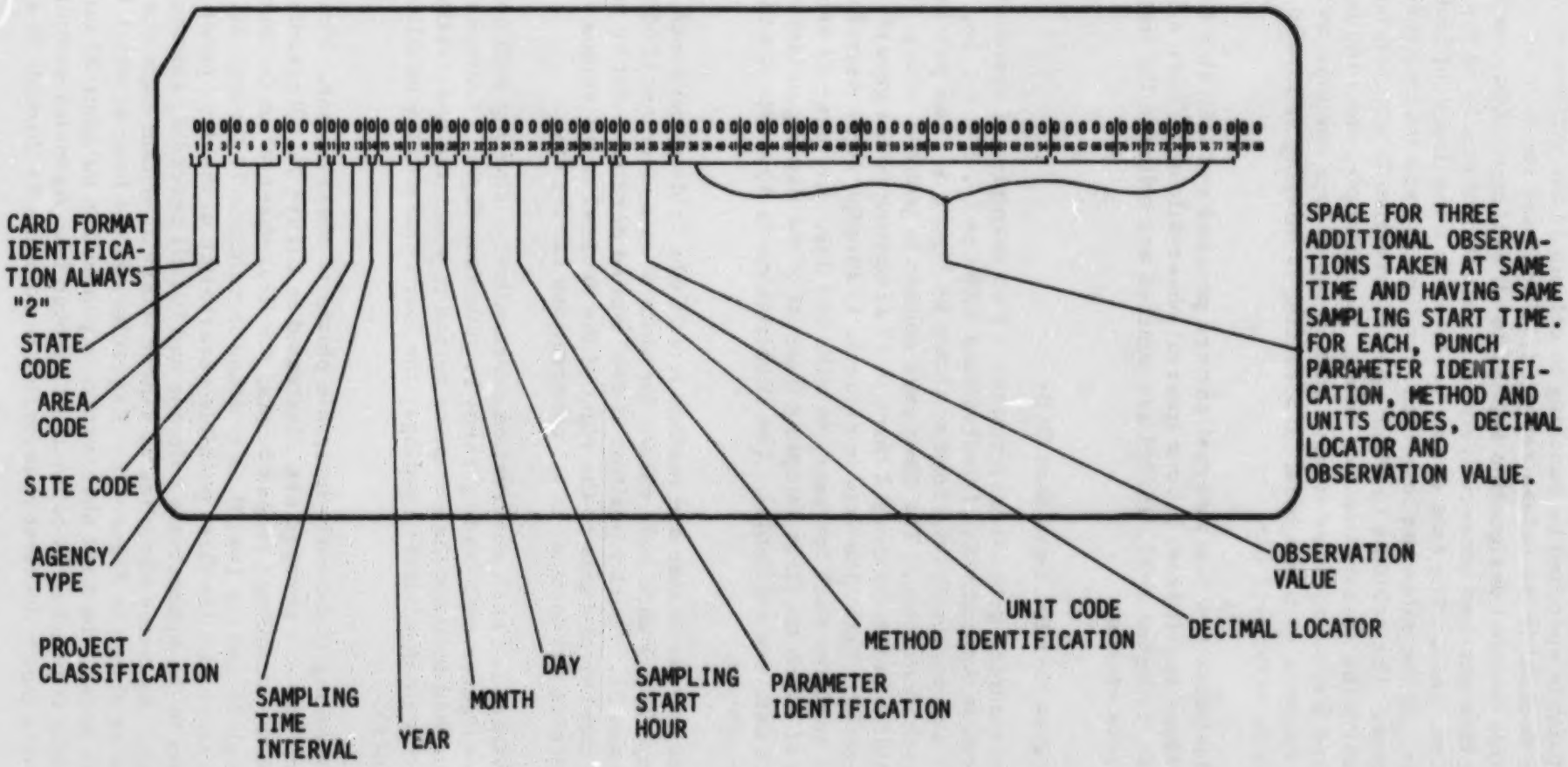


Figure 5. Card format used for data recorded at 24-hour or longer intervals.

implied decimal point is specified by punching in column 32 the number of digits to the right of the decimal in the values reported. Each card can hold 12 observation values, and this decimal designation applies to all readings. Observations for each half day are then punched consecutively starting with column 33 on the appropriate a.m. or p.m. card. The data can be recorded to a maximum of four significant figures. If the value has less than four figures, add leading zeros to provide four figures. The columns in which each observation is punched indicate the hour during which the observation was taken. For example, the sample taken between 7 a.m. and 8 a.m. appears in the eighth group of four columns on the a.m. card, that is, in columns 61 through 64. If a reading is missing or invalid, leave the columns blank or enter four 9's.

Figure 6 illustrates how the observations are punched and indicates those columns whose values result from 1-hour interval observations. Those columns headed with an "s" (Columns 2-13, 15-20) are punched according to the instructions given in the previous section.

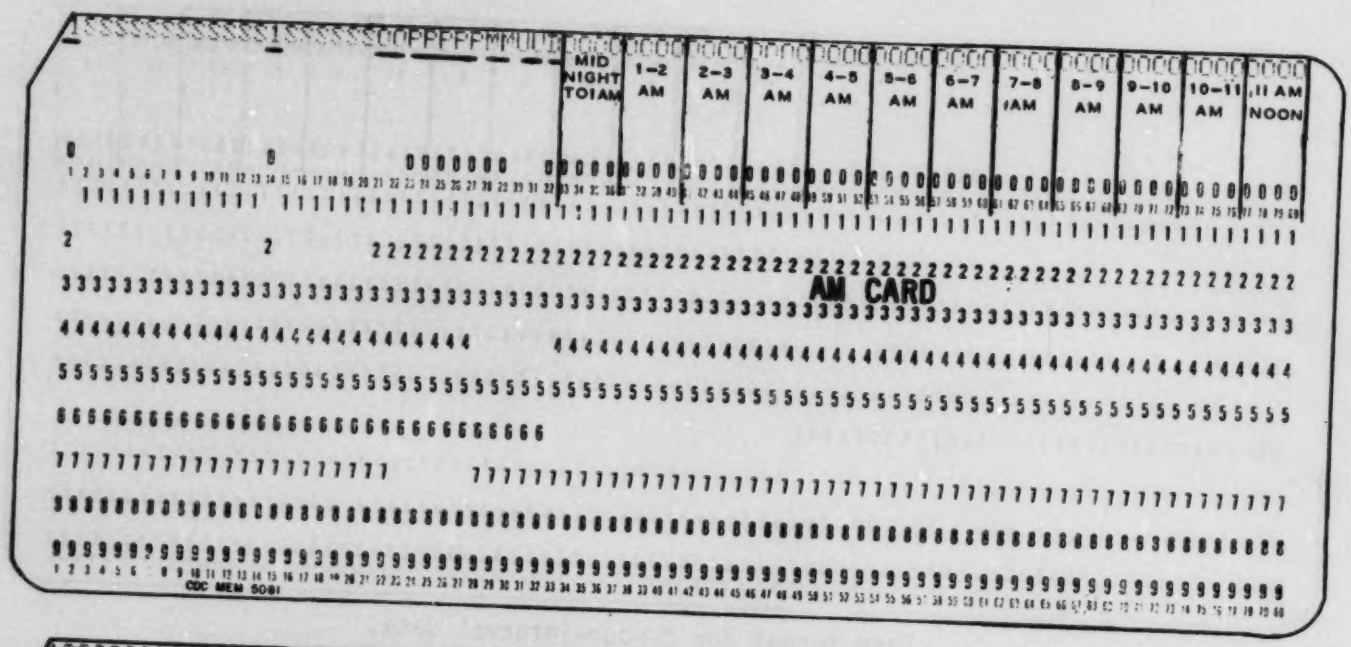
Intervals Longer than Hourly, Less than Daily

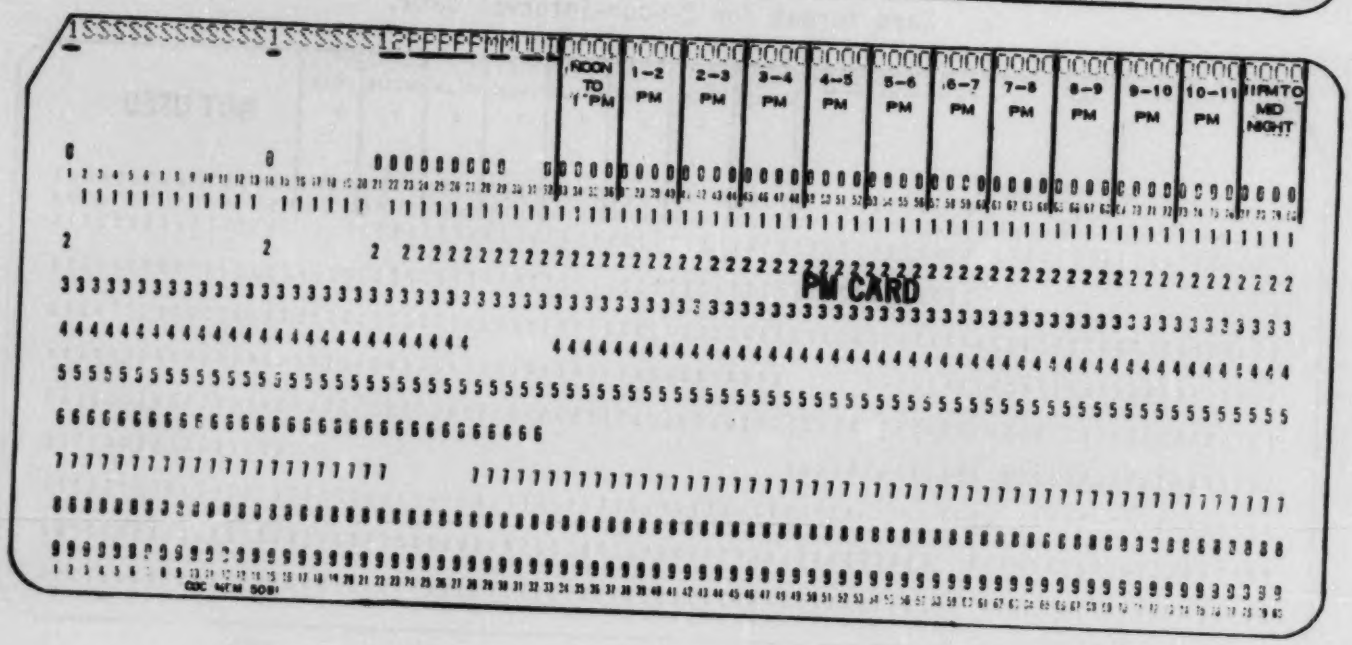
This section contains general instructions. Two examples of hypothetical situations are given in Appendix E. Observations taken at 2-, 3-, 4-, 6-, 8-, or 12-hour intervals are punched in the format shown in Figure 4. One punched card is required for each day's data. The first card column is punched with a "1" to indicate this format. Punch columns 2 through 13 according to the general instructions given on page 15. Enter the date in columns 15 through 20 as described earlier, with two numbers each for year, month, and day. Columns 21 and 22 indicate the hour at which the first sampling interval of the day began (not the first interval for which data are available). The illustrations in Appendix E will clarify details of this procedure.

Obtain codes for parameter and method from Code Table 4 and punch in columns 23 through 29. Punch code values for units of measurement (Code Table 5) in columns 30 and 31. Specify position of the implied decimal point by punching in column 32 the number of digits to the right of the decimal in the values reported. This one value locates the decimal in all observations for the day.

Punch the values for each observation consecutively, starting with column 33, in four significant figures with leading zeros to yield four digits. Columns in which the values are punched indicate the sampling period in which the observation was taken. For missing or invalidated readings, the four columns may be filled in with 9's or left blank.

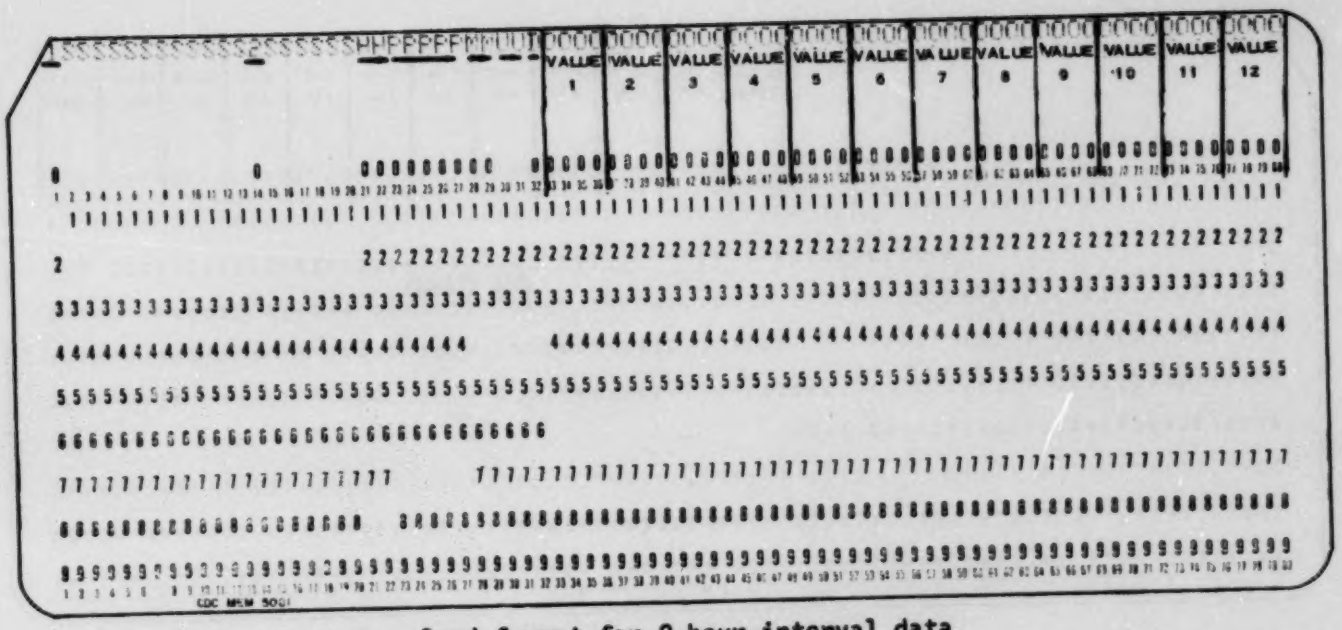
Agencies sampling at 2-hourly intervals obtain 12 observations. Positions in which the observations are punched are illustrated in Figure 7a. Regardless of the hour at which the sampling process began, the first observation is punched in columns 33 through 36 and the twelfth is punched in columns 77 through 80. If sampling did not start with the first possible observation of the day, punch the first actual observation in the appropriate columns and fill all preceding columns with 9's or leave blank. Agencies sampling at 3-hour intervals obtain eight observations per day, punched as shown in Figure 7b. Regardless of the hour at which the sampling process began, the first observation is punched in columns 33 through 36 and the last (eighth) is punched in columns 61 through 64. Agencies sampling at six 4-hour intervals punch the first observation in columns 33 through 36 and the



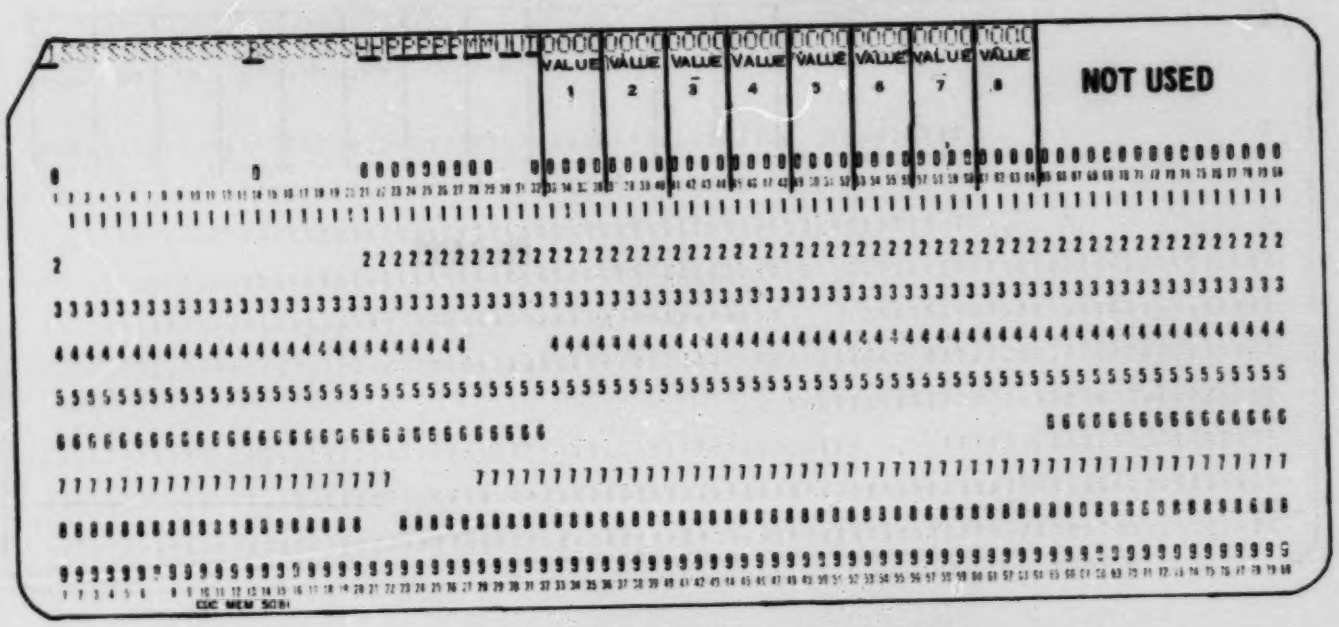


NOTE: S - punched according to general instructions
P - parameter code
M - method code
U - units code
D - number of digits to the right of the decimal
O - observations

Figure 6. Card format for 1-hour interval data.



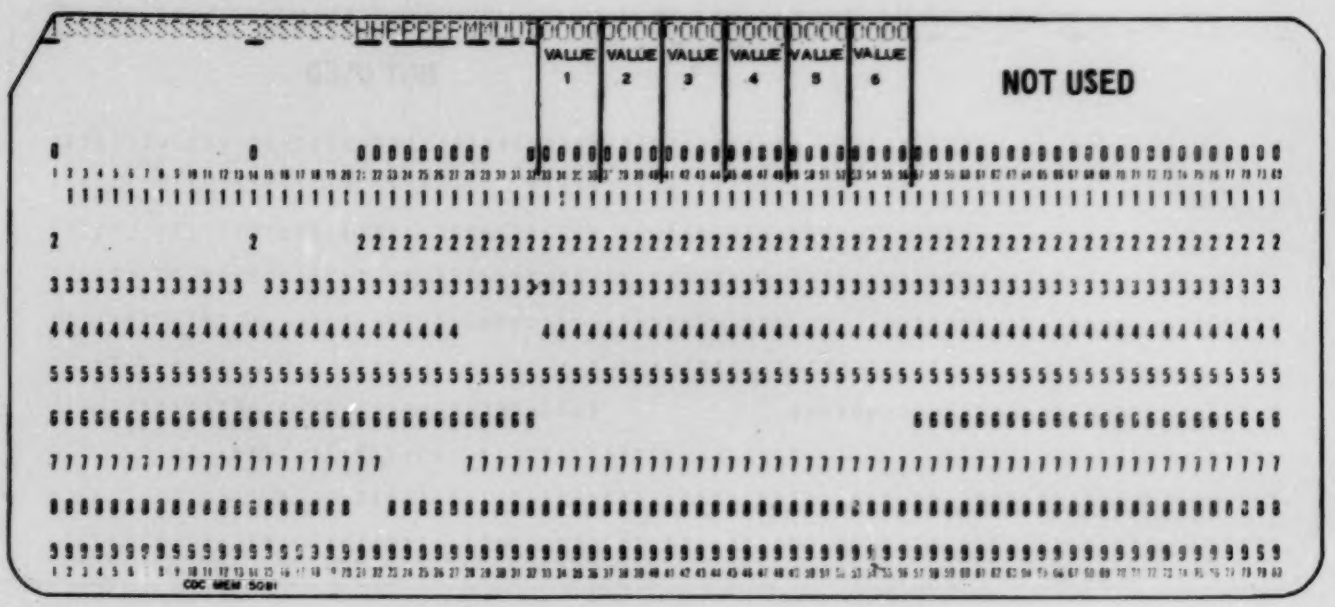
a. Card format for 2-hour-interval data.



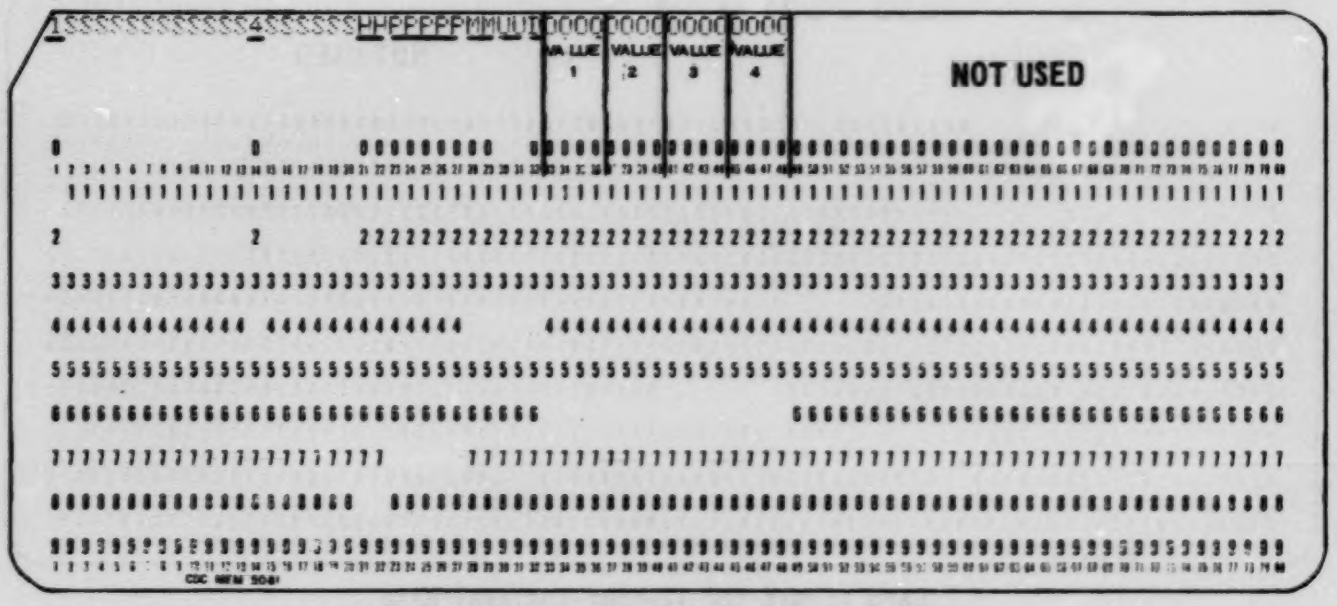
b. Card format for 3-hour-interval data.

NOTE: S - punched according to general instructions
H - start hour
P - parameter code
M - method code
U - units code
D - number of digits to the right of the decimal
O - observations

Figure 7. Various-interval values (less than 24-hour) punched on cards.



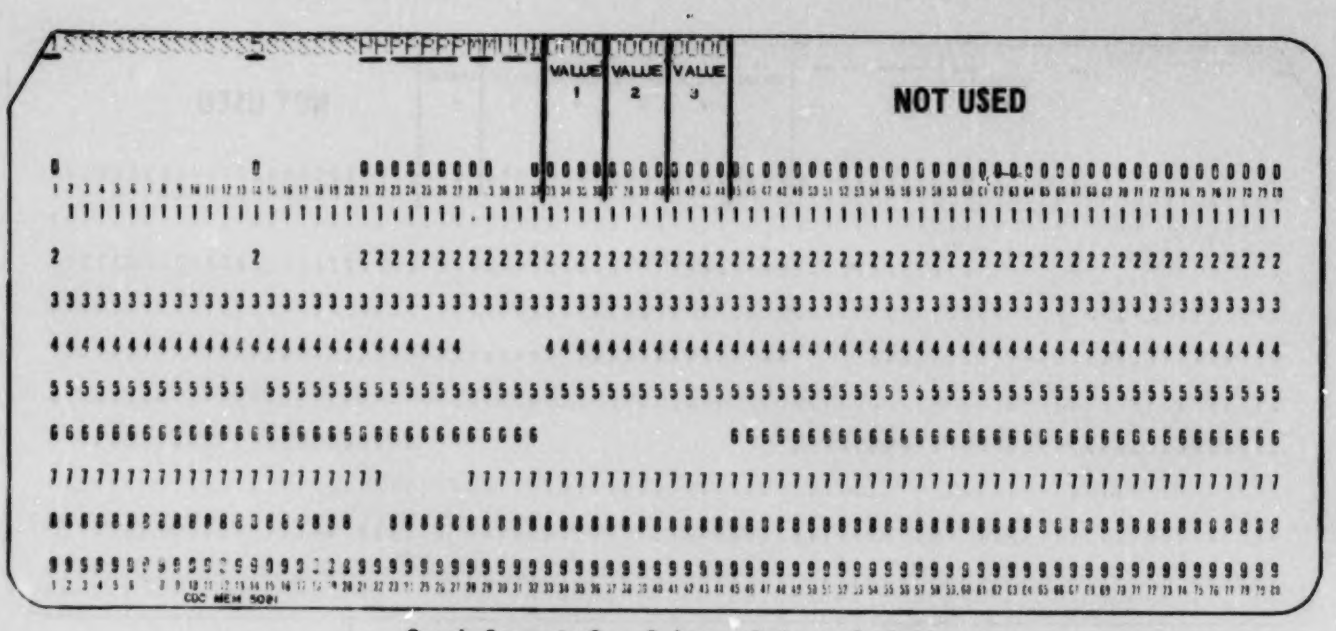
c. Card format for 4-hour-interval data.



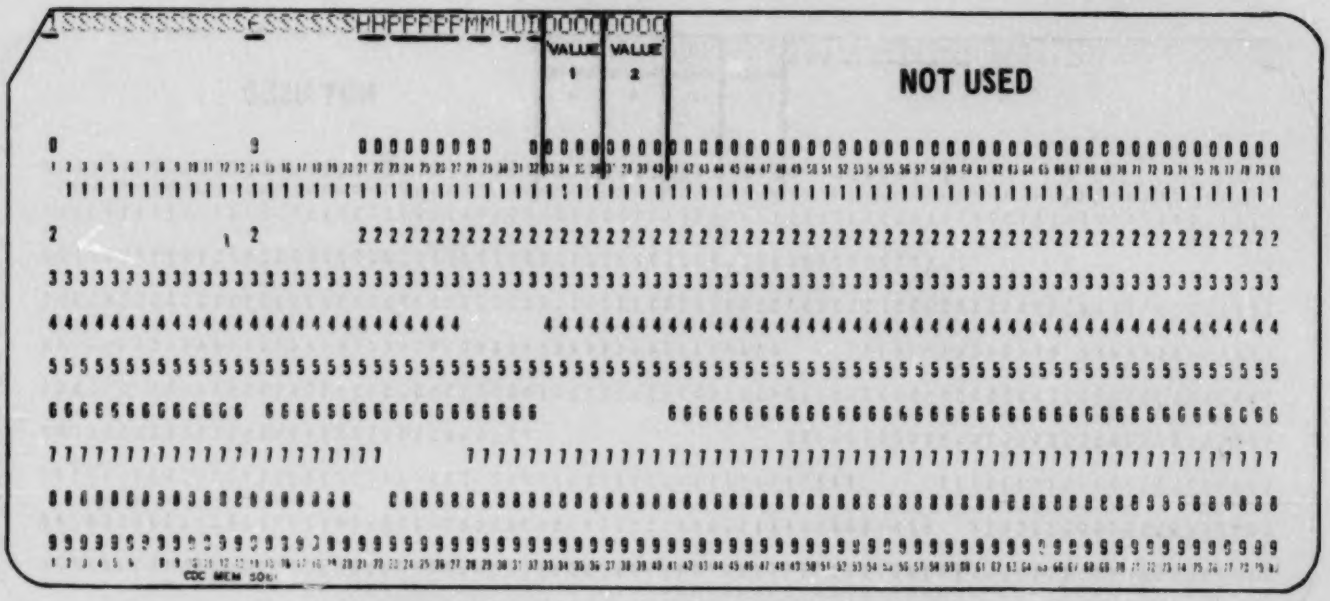
d. Card format for 6-hour-interval data.

NOTE: S - punched according to general instructions
H - start hour
P - parameter code
M - method code
U - units code
D - number of digits to the right of the decimal
O - observations

Figure 7 (continued). Various-interval values (less than 24-hour) punched on cards.



e. Card format for 8-hour-interval data.



f. Card format for 12-hour-interval data.

NOTE: S - punched according to general instructions
P - parameter code
M - method code
U - units code
D - number of digits to the right of the decimal
O - observations

Figure 7 (continued). Various-interval values (less than 24-hour) punched on cards.

sixth in columns 53 through 56 (Figure 7c).

The first of four 6-hour-interval observations is punched in columns 33 through 36 and the last (fourth) is punched in columns 45 through 48 (Figure 7d). The first of three 8-hour-interval observations is punched in columns 33 through 36 and the last (third) in columns 41 through 44 (Figure 7e). The first of two 12-hour-interval observations is punched in columns 33 through 36 and the last (second) in columns 37 through 40 (Figure 7f).

Daily or Longer Intervals

Observations taken at daily or longer intervals are punched according to the general instructions in the format shown in Figure 5.

Data obtained over daily or longer intervals are punched on a card identified by a "2" punched in the first card column. Punch columns 2 through 13 according to the general instructions on page 15. Punch column 14 with the code value obtained from Code Table 3 to indicate the sampling interval (see Figures 8a, b, c, d, as appropriate). In columns 15 and 16 enter the last two digits of the year in which sampling occurred. Punch the month number in columns 17 and 18. Columns 19 and 20 identify the day of the month on which the sample was taken if the interval is daily (see Figure 8a), or the day of the month on which the sampling was begun if the interval is longer than once per day (see Figures 8b, c, d,). For daily-interval data, punch columns 21 and 22 to indicate the hours of the day during which sampling was begun. For longer-interval data, these columns may be left blank.

Obtain the numerical code for the parameter measured from Code Table 4 and punch in columns 23 through 27. Punch the method code (values from Code Table 4) in columns 28 and 29. Find the code indicating the units in which the observation is recorded in Code Table 5 and punch in columns 30 and 31. Since decimal points are not punched on SAROAD formatted cards, specify the position of the implied decimal point by punching in column 32 the number of digits that are to the right of the decimal in the recorded values.

Data can be recorded to four significant figures. If the value has less than four figures, add as many leading zeros as are necessary to give four digits, which are punched in columns 33 through 36.

Observations on up to three additional parameters collected at the same site, at the same time, with a sampling process having the same beginning time may be punched in the remaining columns of this card. For each additional observation, the parameter identification code, the method code, units code, position of the implied decimal point, and value are punched in the same groupings and column widths specified above. If more than four parameters are observed at the same site and at the same time, punch additional cards as required.

Converting Data from SAROAD Forms to Punched Cards

When agencies using the SAROAD forms for storing data acquire an automated data processing capability, punched cards can be produced readily. The SAROAD forms were designed to require a minimum of keypunch instructions. On the form, the numbers below columns in which observations, dates, and parameter

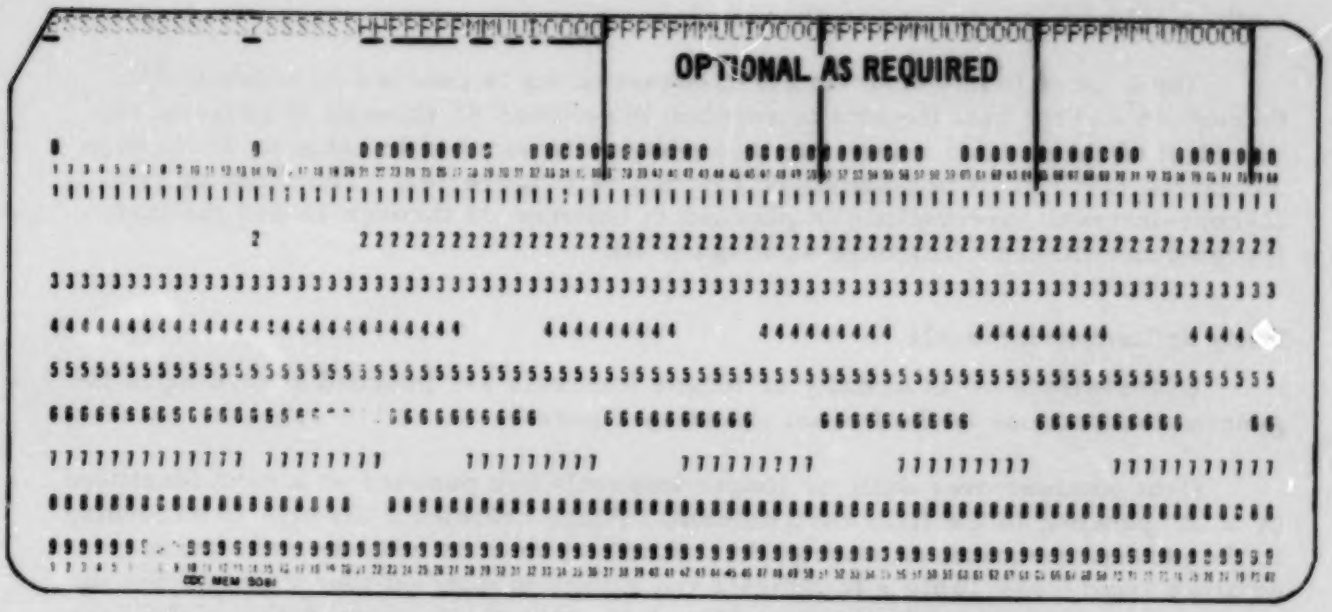


Figure 8a. Card format for daily-interval data.

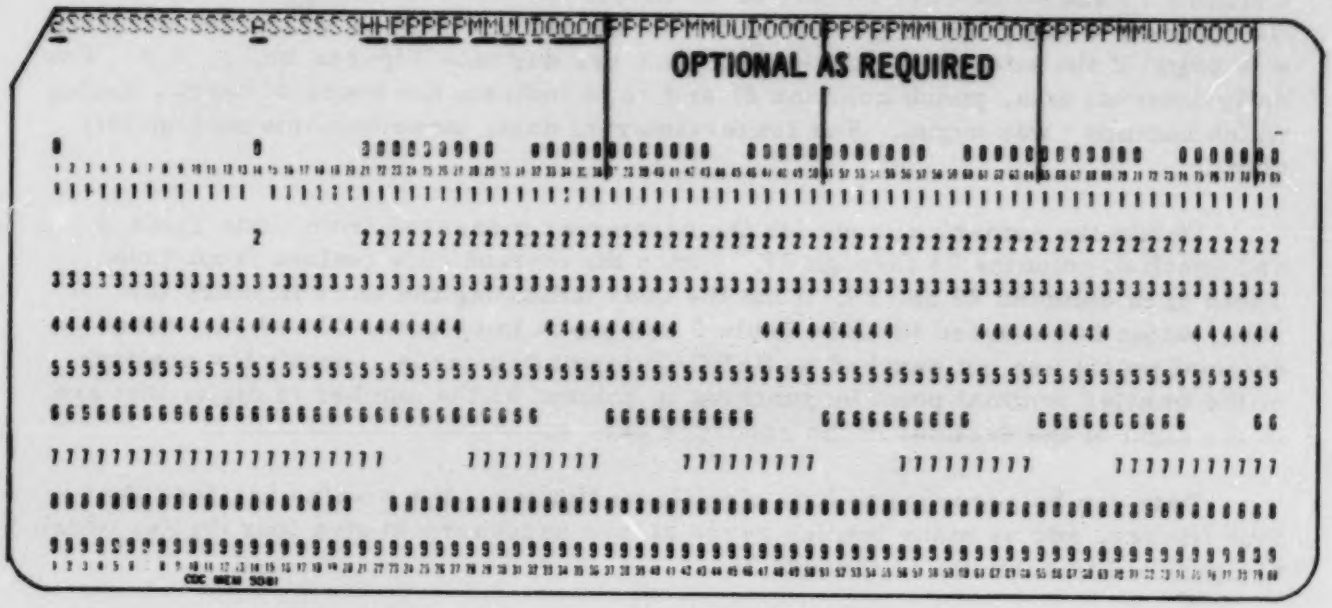


Figure 8b. Card format for weekly-interval data.

NOTE: S - punched according to general instructions
H - start hour
P - parameter code
M - method code
U - units code
D - number of digits to the right of the decimal
O - observations

Figure 8. Card format.

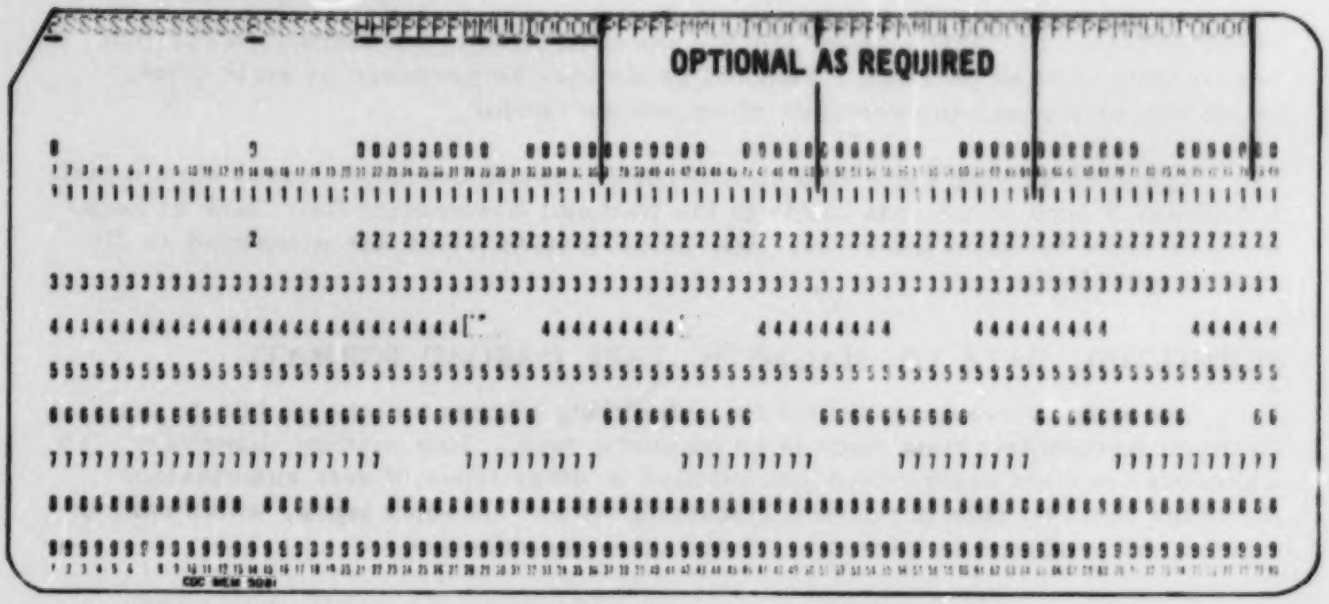


Figure 8c. Card format for monthly-interval data.

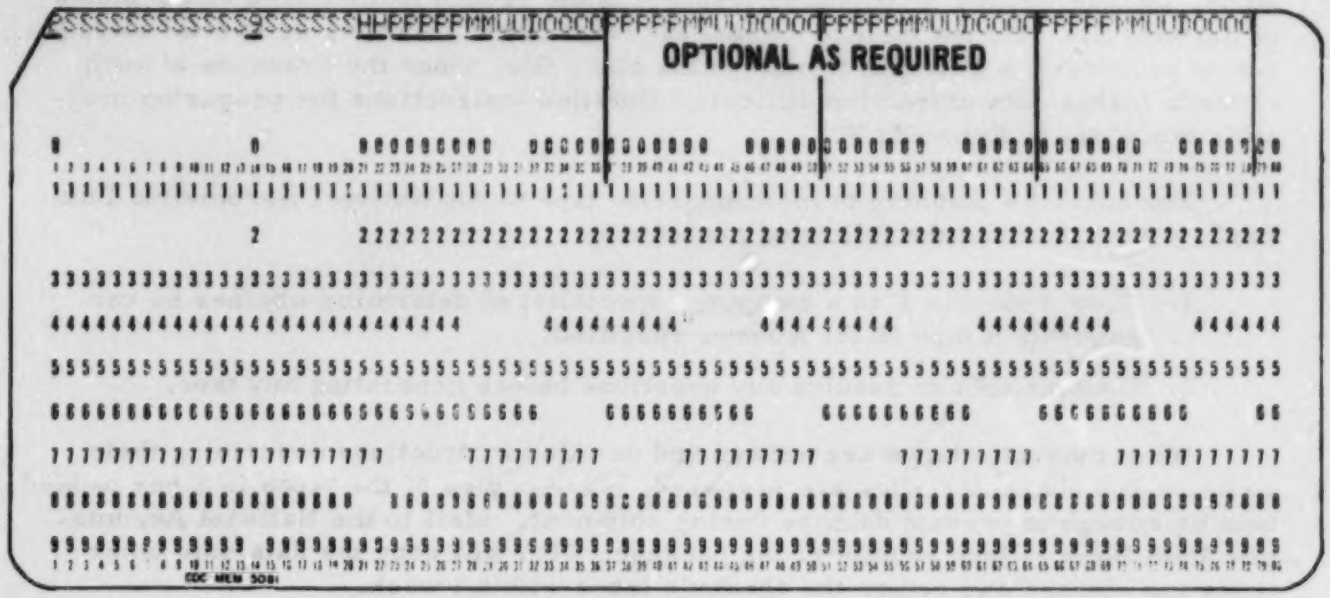


Figure 8d. Card format for quarterly-interval data.

NOTE: S - punched according to general instructions
P - parameter code
M - method code
U - units code
D - number of digits to the right of the decimal
O - observations

Figure 8 (continued). Card format.

identification codes are written indicate the card column positions in which that information is to be punched. Punched cards may be prepared by an in-house capability or through the services of an outside vendor.

Agencies storing their data on SAROAD-formatted punched cards should mail a duplicated deck of the data cards to the National Aerometric Data Bank at quarterly or more frequent intervals. Site Identification forms are submitted as discussed previously.

SUBMITTING DATA ON MAGNETIC TAPE (SAROAD FORMAT)

The most convenient method for submitting a large volume of data to the National Aerometric Data Bank is on magnetic tape. This method, however, also introduces certain restrictions not entailed in other types of data submission. Agencies must be able to provide IBM-compatible* 1/2-inch tapes, which can, of course, be produced by computers from other manufacturers.

Agencies capable of providing this type of tape are encouraged to submit SAROAD-formatted card image records of their data. When submitting magnetic tapes, include written instructions identifying the record format used and contents of the data file. Do not include extraneous data, blank records, or special identification records of a different format in the same file, since the presence of such records makes data extraction difficult. Detailed instructions for preparing magnetic tapes are in Appendix F.

For agencies planning to mail magnetic tape to the National Aerometric Data Bank, we recommend the following procedure:

1. Show Appendix F to a computer specialist to determine whether he can generate a tape in the manner specified.
2. Contact EPA to resolve any questions before generating any tape.

When magnetic tapes are written and detailed instructions describing their contents and characteristics are prepared, place copies of the tapes in a box padded heavily enough to prevent damage during shipment. Mail to the National Aerometric Data Bank. Upon receipt of the package, EPA will copy the data onto other tapes for storage and return the agency's tapes within 1 week.

---

*Use of the term 'IBM-compatible' does not constitute an endorsement of any manufacturer but reflects data processing capabilities of EPA.

CONVERTING EXISTING DATA TO SAROAD FORMAT

A considerable amount of past data are available in a variety of non-SAROAD formats. Also, many agencies who have developed their own format systems over the years may believe that wholesale conversion of data to the SAROAD format now involves too great a disruption in the established agency routine. These agencies may, however, maintain data in their own format for internal use and also provide copies in SAROAD format for entry into the National Aerometric Data Bank and for exchange with other organizations. In this case, the agency will require a system for converting from their format to SAROAD format. The following general guidelines will be helpful for developing conversion programs.

There are three possible conversions:

1. Cards to SAROAD cards.
2. Cards to SAROAD tape.
3. Tape to SAROAD tape.

If agency data are now on forms, key punch these data onto cards. It is imperative that a consistent format be used in preparing the punched cards. In organizing this format a computer specialist should be consulted to insure machine readability.

To establish an effective format conversion program, an agency proceeds through the following steps:

1. Obtain from EPA unique SAROAD identification numbers for all agency sampling sites.
2. Determine the SAROAD codes for parameters, methods, time intervals, units of observation, agency type, and project classification.
3. Determine whether punched cards or magnetic tape will be used for reporting data.
4. Identify the SAROAD format that corresponds to the frequency of sampling.
5. Develop tables that show correspondence between existing agency codes and the appropriate SAROAD codes.
6. Develop a simple program to convert agency format into SAROAD format. This conversion involves the following steps:
   a. Read into computer memory a punched card or a record from magnetic tape.
   b. Based on the type of data, decide on the format to be used for the output.
   c. By use of the tables, determine the appropriate SAROAD codes corresponding to the agency codes. These codes include FORM NUMBER, STATE, AREA, SITE, AGENCY, PROJECT, TIME INTERVAL,

PARAMETER(s), METHOD(s), UNIT(s), AND DP(s). Transfer these codes into the output area.

d. Take the OBSERVATIONS, YEAR, MONTH, DAY, and START HOUR from the input area, rearrange them if necessary to fit into SAROAD format, and transfer them to the output area (rearrangement may mean changing to Year, Month, Day from Month, Day, Year or adding leading zeros to observations).

e. Punch an output card or write an output record onto magnetic tape.

Follow steps a through e until all the data are converted.

SUBMITTING PAST DATA NOT IN SAROAD FORMAT

A few control agencies may be unable to convert to the SAROAD format immediately. During the interval before conversion to SAROAD is accomplished, the data from those agencies will be needed for the National Aerometric Data Bank. When agencies supply data in a format other than SAROAD, EPA needs information on monitoring sites, methods, and data formats to make the data accessible and understandable. Appendix G discusses how to prepare these data. Experience in receiving data in various non-SAROAD formats has indicated some common pitfalls:

1. Inadequate definition of codes used to indicate method of measurement, pollutant or other parameter, and sampling site.
2. Failure to specify the units in which observations are recorded.
3. Incomplete specification of the location of implied decimal points when decimal points are not explicitly punched in the data records.
4. Omission of the time interval of measurements.
5. Identification by means of colors, handwriting on cards, or tab cards. These "identifications" are not machine-readable and can seriously impede the handling and processing of data. One problem involved in differentiating by color is that information cannot be photocopied.

Note again that providing this information should not be regarded as an extra burden on agency personnel because this information must be available in a systematic form for the agency's own use if data are to be readily accessible and understandable.

Code Table 1. AGENCY TYPE

A	EPA group responsible for atmospheric surveillance
B	EPA group responsible for meteorological activity
C	EPA group responsible for effects research
D	EPA group responsible for atmospheric research
E	EPA group responsible for abatement activity
F	State agency
G	County agency
H	City agency
I	District agency
J	Private
K	Institution (university, college, etc.)
L	Military
M	International agency
N	Other Federal nonmilitary agencies
O-Y	Open for future expansion
Z	Other

Code Table 2. PROJECT CLASSIFICATIONS

Long-term surveillance codes	
01	Population-oriented surveillance
02	Source-oriented ambient surveillance
03	Background surveillance
Short-term surveillance codes	
04	Complaint investigation
05	Special studies
06	Episode monitoring
07-99	For Federal networks and future expansion

Code Table 3. TIME INTERVAL

Code	Data observed over a period of:
1	1 hour
2	2 hours
3	4 hours
4	6 hours
5	8 hours
6	12 hours
7	24 hours
8	1 month
9	3 months
A	1 week
B	3 hours
C	Composite data
D-Z	For future expansion

Code Table 4. COMMON PARAMETERS AND METHODS[a]

Parameter	Parameter code	Method of collection and analysis		
		Collection	Analysis	Method code
Suspended particulate				
Total suspended particulate	11101	Hi-Vol	Gravimetric	91
Benzene-soluble organics	11103	Hi-Vol	Benzene extraction-Soxhlet	91
Soiling index (COH/1000LF)	11201	Tape sampler	Transmittance	81
Soiling index (RUD/10000LF)	11202	Tape sampler	Reflectance	91
Light scatter	11203	Integrating nephelometer		11
Radioactivity (gross beta)	11302	Hi-Vol	Proportional counter	91
Antimony	12102	Hi-Vol	Emission spectra	91
Antimony	12102	Hi-Vol	Atomic Absorption	92
Arsenic	12103	Hi-Vol	Atomic absorption	92
Beryllium	12105	Hi-Vol	Atomic absorption	92
Bismuth	12106	Hi-Vol	Atomic absorption	92
Barium	12107	Hi-Vol	Atomic absorption	92
Boron	12108	Hi-Vol	Atomic absorption	92
Cadmium	12110	Hi-Vol	Atomic absorption	92
Chromium	12112	Hi-Vol	Atomic absorption	92
Cobalt	12113	Hi-Vol	Atomic absorption	92
Copper	12114	Hi-Vol	Atomic absorption	92
Iron	12126	Hi-Vol	Atomic absorption	92
Lead	12128	Hi-Vol	Atomic absorption	92
Manganese	12132	Hi-Vol	Atomic absorption	92
Molybdenum	12134	Hi-Vol	Atomic absorption	92
Nickel	12136	Hi-Vol	Atomic absorption	92
Mercury	12142	Hi-Vol	Atomic absorption	92
Selenium	12154	Hi-Vol	Atomic absorption	92
Tin	12160	Hi-Vol	Atomic absorption	92
Vanadium	12164	Hi-Vol	Atomic absorption	92
Zinc	12167	Hi-Vol	Atomic absorption	92
Fluoride	12202	Hi-Vol	Specific ion electrode	91
Ammonium	12301	Hi-Vol	Nessler	91
Ammonium	12301	Hi-Vol	Sodium phenolate	92
Nitrate	12306	Hi-Vol	2-4 xylenol	91
Nitrate	12306	Hi-Vol	Reduction-Diazo coupling	92
Sulfate	12403	Hi-Vol	Colorimetric	91

Code Table 4 (continued). COMMON PARAMETERS AND METHODS[a]

Parameter	Parameter code	Method of collection and analysis		
		Collection	Analysis	Method code
Suspended particulate (continued)				
Sulfate	12403	Hi-Vol	Turbidimetric	92
Benzo(A)pyrene	17242	Hi-Vol	Thin-layer chromatography	91
Settled particulate (dustfall)				
Total particulates	21101	Bucket; jar	Gravimetric	81
Sulfation rate				
Sulfation rate	42410	Lead plate	Gravimetric	71
Sulfation rate	42410	Lead plate	Potassium carbonate	74
Sulfation rate	42410	Lead candle	Gravimetric	91
Sulfation rate	42410	Lead candle	Titrimetric	92
Gases				
Carbon monoxide	42101	Instrumental	Nondispersive infrared	11
Sulfur dioxide	42401	Instrumental	West-Gaeke colorimetric	11
Sulfur dioxide	42401	Instrumental	Conductometric	13
Sulfur dioxide	42401	Instrumental	Coulometric	14
Sulfur dioxide	42401	Gas bubbler	West-Gaeke	91
Sulfur dioxide	42401	Gas bubbler	Conductometric	93
Nitric oxide	42601	Instrumental	Colorimetric	11
Nitrogen dioxide	42602	Instrumental	Colorimetric	11
Nitrogen dioxide	42602	Gas bubbler	Jacobs-Hochheiser	91
Oxides of nitrogen	42603	Instrumental	Colorimetric	11
Ammonia	42604	Gas bubbler	Nessler	91
Ammonia	42604	Gas bubbler	Sodium phenolate	92
Aldehyde	43501	Instrumental	Colorimetric	11
Total oxidants	44101	Instrumental	Alkaline KI	11
Total oxidants	44101	Instrumental	Neutral KI colorimetric	14
Total oxidants	44101	Gas bubbler	Alkaline KI	81
Total oxidants	44101	Gas bubbler	Phenolphthalein	51
Ozone	44201	Instrumental	Chemiluminescence	11
Methane	43201	Instrumental	Flame ionization	11
Hydrocarbon	43101	Instrumental	Flame ionization	11

Code Table 4 (continued). COMMON PARAMETERS AND METHODS[a]

Parameter	Parameter code	Method of collection and analysis		
		Collection	Analysis	Method code
Meteorological observations				
Wind speed	61101	Instrumental		11
Wind direction	61102	Instrumental		11

[a]This list covers the methods used by about 95 percent of the currently operating control agencies. If a parameter or method is not listed here, a code can be assigned upon evaluation of the method by EPA.

Code Table 5. UNITS

Code number	Units
01	micrograms/cubic meter (25° C, 1013 millibars)
02	micrograms/cubic meter (0° C, 1013 millibars)
03	nanograms/cubic meter (25° C, 1013 millibars)
04	nanograms/cubic meter (0° C, 1013 millibars)
05	milligrams/cubic meter (25° C, 1013 millibars)
06	milligrams/cubic meter (0° C, 1013 millibars)
07	parts per million (volume/volume)
08	parts per billion (volume/volume)
09	COHS/1000 linear feet
10	RUDS/10,000 linear feet
11	meters/second
12	miles/hour
13	knots
14	degrees
20	microns
30	picocuries/cubic meter
31	microcuries/cubic meter
32	picocuries/square meter
33	microcuries/square meter
34	picocuries/cubic centimeter
35	picocuries/gram
50	number of threshold levels
70	milligrams F/100 square centimeters-day
80	milligrams SO_3/100 square centimeters-day
81	micrograms SO_2/square meter-day
90	tons/square mile-month[a]
91	milligrams/square centimeter-month[a]
92	micrograms/cubic meter-month[a]
98	milligrams SO_4^{2-}/square centimeters-30 days
99	milligrams/square centimeters-30 days

[a]On a calendar-month basis.

Code Table 6. MATRIX FOR START-HOUR VALUES

First possible hour sampling could occur	Sampling interval between observations, hours					
	2	3	4	6	8	12
Midnight	00	00	00	00	00	00
1 AM	01	01	01	01	01	01
2 AM		02	02	02	02	02
3 AM			03	03	03	03
4 AM				04	04	04
5 AM		THESE		05	05	05
6 AM		CANNOT			06	06
7 AM		OCCUR			07	07
8 AM						08
9 AM						09
10 AM						10
11 AM						11

APPENDIX A.
DEVELOPING A MINIMAL COMPUTERIZED DATA PROCESSING SYSTEM

Although the desire of an agency to avoid becoming unnecessarily involved with computers is commendable, a well-designed computer-based system can be a valuable tool in certain circumstances. EPA estimates that when more than one man-week per month of professional time, in addition to clerical time, is expended in making calculations, manipulating data, and preparing data summaries and analyses, a control agency should develop a minimal computer-based data processing system. Also if more than one employee spends full time processing data, machine processing is justifiable. If a control agency intends to use elaborate data acquisition systems such as those that employ telemetry, or to use advanced analysis techniques such as modeling, it is essential that they have developed a workable computerized data handling system.

An agency should first explore whether a system and computer time are available from another control agency. It may be possible to pool resources with other agencies in developing such a system. In any case, the steps required to establish such a system are as follows:

1. Define the manual system you are currently using in great detail. Define specific problems in the system and how a computer can be used to alleviate them. If you are unable to accomplish this first step, you are not in a position to develop a computer-based system.
2. Define each separate process in the new system, taking care to demonstrate the man-machine interaction (e.g., after defining a computerized validation program, one must define how the output of the program will be used and what other manual checks the data will receive).
3. Arrange for keypunch services and computer time.
   a. Keypunch: Take the raw aerometric data to a university, a government agency, or a private computer service bureau for keypunching. Request that the data be keypunched in the SAROAD formats as explained in this manual. Arrange to have future keypunching done at regular intervals.
   b. Computer: Arrange for computer time with a university, a government agency, or a private computer service bureau. This must be done in advance, for two reasons. First, a computer is needed on which time will be available for the future. Second, computer programs are for a specific computer (manufacturer and model) and to use a program on another computer often requires modification.
4. Arrange for a programmer in a government agency, a university, or a private computer service bureau to write a validation program that will list the data appearing on the data cards, "flagging" or marking off suspected readings if:

a. They are above a preset upper limit.

b. They are below a preset lower limit.

c. They represent greater than a preset percentage change from a preceding reading.

5. Arrange for a programmer to write a program that will:

a. Pick off high readings and time of occurrence.

b. Calculate arithmetic and geometric means and their respective standard deviations.

c. Print out data (or summaries of data) on computer paper in a format that can be photocopied and immediately incorporated in periodic reports.

If an agency's capability is developed along these lines, professional personnel need only insure that raw data are accessible to the keypunch operators, review computer output to validate the data, and interpret the analyzed data, which can then be immediately incorporated into the agency's periodic reports. No other involvement with routine data handling will be necessary, and valuable professional manpower will be freed for more important work.

Hazards to be avoided in introducing a computerized data-handling system include:

1. Attempting to develop a more elaborate system than is required. By far the best approach is to begin with the simplest possible system, make it operational, and then add increased capability later if it becomes necessary. As a general rule, KEEP IT SIMPLE!

2. Providing insufficient direction to computer personnel. If all necessary information is not perfectly clear to computer personnel, time and money will be wasted on development of ineffective programs.

3. Not requiring sufficient documentation. If the programmer who establishes the formats and writes the programs should no longer be available to the agency, the entire system might be rendered useless. For this reason, the formats and programs used in developing a computerized data processing system must be clearly and completely described. (Note that the standardized SAROAD format is completely documented in this manual.) Also, an agency should demand that any program be completely documented so that it is accessible to any programmer familiar with the computer language in which the program was written.

If an agency avoids these pitfalls, certain advantages will result from a simple computer-based data processing system: availability of agency data will be increased, and valuable manpower will become less involved with routine data handling and calculation. These advantages are sufficient to justify careful development of a basic computerized system.

FURTHER WORD ON DOCUMENTATION

Among the several pitfalls mentioned, lack of documentation is one that an agency can avoid by careful planning. As a further guide to establishing a minimal computerized system, we outline here a basic approach to system documentation.

System Description and Flow

The first step in documentation involves an overview of the entire system. Verbal description should define the functions of the various computer routines, point out the points of interface between the computer and agency personnel, and show the flow of data. An example of this flow is presented in the diagram shown in Figure A-1. The figure is not intended to represent a complete system, but merely exemplifies a part of a system.

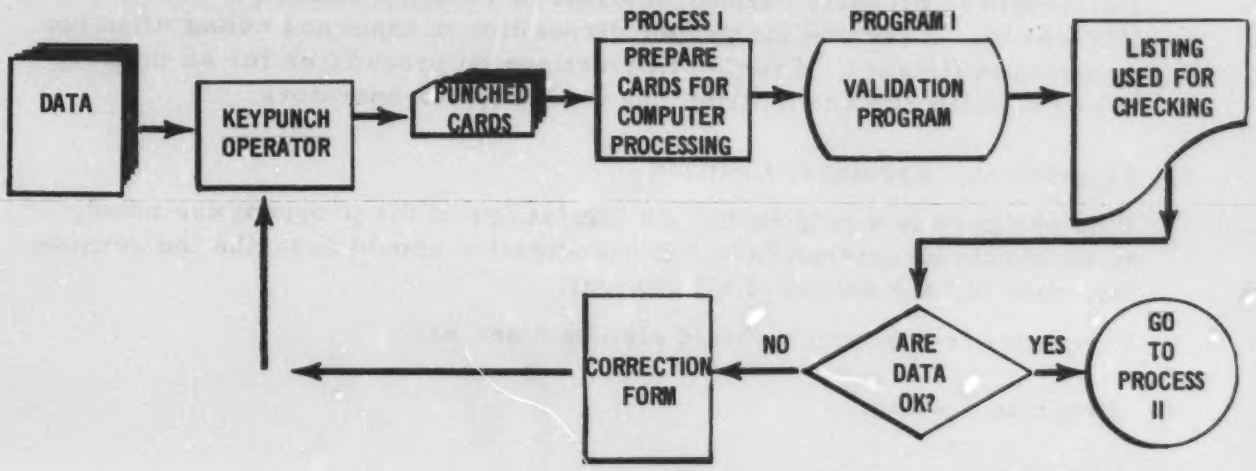


Figure A-1. Portion of typical flow diagram.

Documentation for Individual Computer Programs

Documentation should be complete, so that whoever is responsible for maintaining and altering the computer program in the future will have enough information to do so. Program documentation usually consists of six phases.

1. Program Description

   A full description of the program should include specifications for design of the program (what it does), restrictions and limitations, the procedures used, any formulas included, options available, unique features (if any), and a list of related programs.

2. Flow Chart

   The chart should show the detailed logic of the program and should be complete enough that someone not familiar with the program could make changes.

3. Input/Output Description

   For punched cards the following elements are needed: control card description, data card description, end card description, and card deck

setup. Many computer programs allow options for input or output. The control card defines which option to take.

For magnetic tape or direct access storage devices, such items as file design, sequence, header, or end records, are defined here.

A copy of the output description and format should be presented. The output can be on magnetic tape, punched cards, or printed copy.

4. Operating Instructions

Instructions should be provided for filling out control cards and end cards, keypunching for data, control, and end cards, and preparing the input deck and any file or peripheral devices. Indicate frequency and length of run (length is probably defined in terms of volume, such as 5 minutes/1000 cards). Describe the proper disposition of input and output when the program is finished. Provide instructions on procedures for an interrupted or aborted run and instructions for computer operators.

5. Suggestions, Warnings, Changes

This category is a potpourri. As limitations of the program are noted, these should be entered here. Documentation should describe the reasons for, date of, and nature of all changes.

Frequent error sources should also be noted here.

6. Program Listing.

APPENDIX C.
OPTIONAL FORMS

Two optional forms may be useful in certain circumstances.

The first, shown in Figure C-1, is printed on a 5- by 8-inch card in the same format as the paper Daily Data Form, which can provide data for a month for four or fewer pollutants. Although this form can be used to record any data with a sampling interval of 24 hours or greater, its main advantage is that one form can accommodate 24 parameters. The face of the card provides space to record four parameters; the back provides space for 20 additional parameters. This card can be used for recording values when a Hi-vol filter is analyzed for a number of subfractions. It is easy to store for later entries if further analyses are to be performed in the future. For this purpose, the form provides space to record a filter number (for cross-reference) and air volume.

The second of the optional forms, shown in Figure C-2, is known as the SAROAD Multiple-Station Form. It was developed for recording data obtained at a sampling interval of 24 hours or greater for one parameter for up to 12 sampling sites. A control agency operating a network of 30 dustfall buckets, for example, could use three of these forms, as opposed to using 30 of either of the Form 2's to record a month's data.

The Multiple-Station Form provides space for identifying information in the upper left. Those standard items are entered as on the other forms. Note that this form provides room for three sets of four sites, from site address A in the upper left to site address L in the lower left above the third set of blocks.

For entries in the blocks in the upper right, the two-digit state code is assigned by EPA, agency type code is given in Code Table 1, project code in Code Table 2, and time interval code in Code Table 3.

For entries on the next line, the five-digit parameter code and the two-digit method code are given in Code Table 4 and the units code in Code Table 5. The "DP" (decimal point) block indicates the number of places to the right of the decimal point in the value entries. Example: a value of 12.1 $\mu g/m^3$ is entered as 0121 in the value blocks with a DP of 1.

For "Year" record the last two digits of the year in which the observations were taken; list months numerically in two digits from January (01) through December (12): indicate day of month on which sampling occurred or began in two digits, from 01 to 31.

Next come three sets of blocks, each set containing four lines. Each line can accommodate information from one site. In the first two blocks record the hour the sampler started. (For data with a sampling interval greater than 24 hours, these blocks are probably left blank.) In the next seven blocks record the area and

ENVIRONMENTAL PROTECTION AGENCY
National Aerometric Data Bank
P.O. Box 12055
Research Triangle Park, N.C. 27711

2

1 AGENCY

CITY NAME

SITE ADDRESS

PROJECT

TIME INTERVAL

STATE AREA SITE
2 3 4 5 6 7 8 9 10

AGENCY PROJECT TIME YEAR MONTH
11 12 13 14 15 16 17 18

DAY ST HR
19 20 21 22

PARAMETER NAME METHOD UNITS

FILTER NUMBER
PARAMETER CODE METHOD UNITS DP
(23–32)
(37–46)
(51–60)
(65–74)

AIR VOLUME

PARTICULATE WEIGHT
(33–36)
(47–50)
(61–64)
(75–78)

(FRONT)

PARAMETER NAME METHOD UNITS
PARAMETER CODE METHOD UNITS DP
(23–32)
(37–46)
(51–60)
(65–74)
(33–36)
(47–50)
(61–64)
(75–78)

(BACK)

Figure C-1. SAROAD Multiple-Parameter Data Form.

APPENDIX B.
SITE DESCRIPTION DEFINITIONS

The station type is now defined at two levels. The first level defines a sampling site as being in a

- CENTER CITY area
- SUBURBAN area
- RURAL area
- REMOTE area

Note that the CENTER CITY category is intended as definition of the core area of the city, not its incorporated limits.

CENTER CITY OR SUBURBAN

For the second level of coding, it is necessary to estimate the dominating influence within a 1-mile radius on the sampling sites. The following categories are possible:

1. Industrial: Implies product-oriented establishments such as manufacturing concerns, utilities, mining, and graineries.
2. Commercial: Implies service-oriented establishments. A unique traffic pattern into and out of the area would be expected. Retail establishments, shopping centers, gas stations, laundromats, etc., comprise this category.
3. Residential: Because many other areas are also used residentially, select this category only in the absence of a dominating industrial or commercial influence.
4. Mobile: Category for samplers placed in airports, truck or bus terminals, on an expressway cloverleaf, or in similar locations. Samplers placed near parking lots would probably be better categorized as industrial or commercial.

RURAL

For rural sampling sites, three other categories in addition to commercial and industrial are used.

1. Near urban: Category for samplers placed in a rural area, yet close enough to a major urban center to be materially affected by the urban area.
2. Agricultural: Category encompassing orchards, crop raising, cattle and sheep grazing, etc.
3. None of the above.

REMOTE

Remote sites are normally far enough away from any activity to measure geophysical background levels. Any potential influencing source may be described in the space for comments.

Space for comments (132 characters) is provided on the Site Form to help define unique characteristics of the sampling site. This space could be used to mention the category of industry that dominates the area if that is discernable (e.g., lead plant; pulp and paper manufacturing), or to mention other influences (e.g., This site is also influenced by a shopping center 5 blocks away.). Almost every suburban or center city site is affected to some degree by combinations of the categories. Situations that make it difficult to determine a dominating influence can be explained in the comments. It may also be necessary to explain short-term sources such as construction of a new building (e.g., Parking garage built 200 yards to the north from May 1970 - January 1971.). The comments will be sent with other site information in response to requests for data.

ENVIRONMENTAL PROTECTION AGENCY
National Aerometric Data Bank
P.O. Box 12055
Research Triangle Park, N.C. 27711

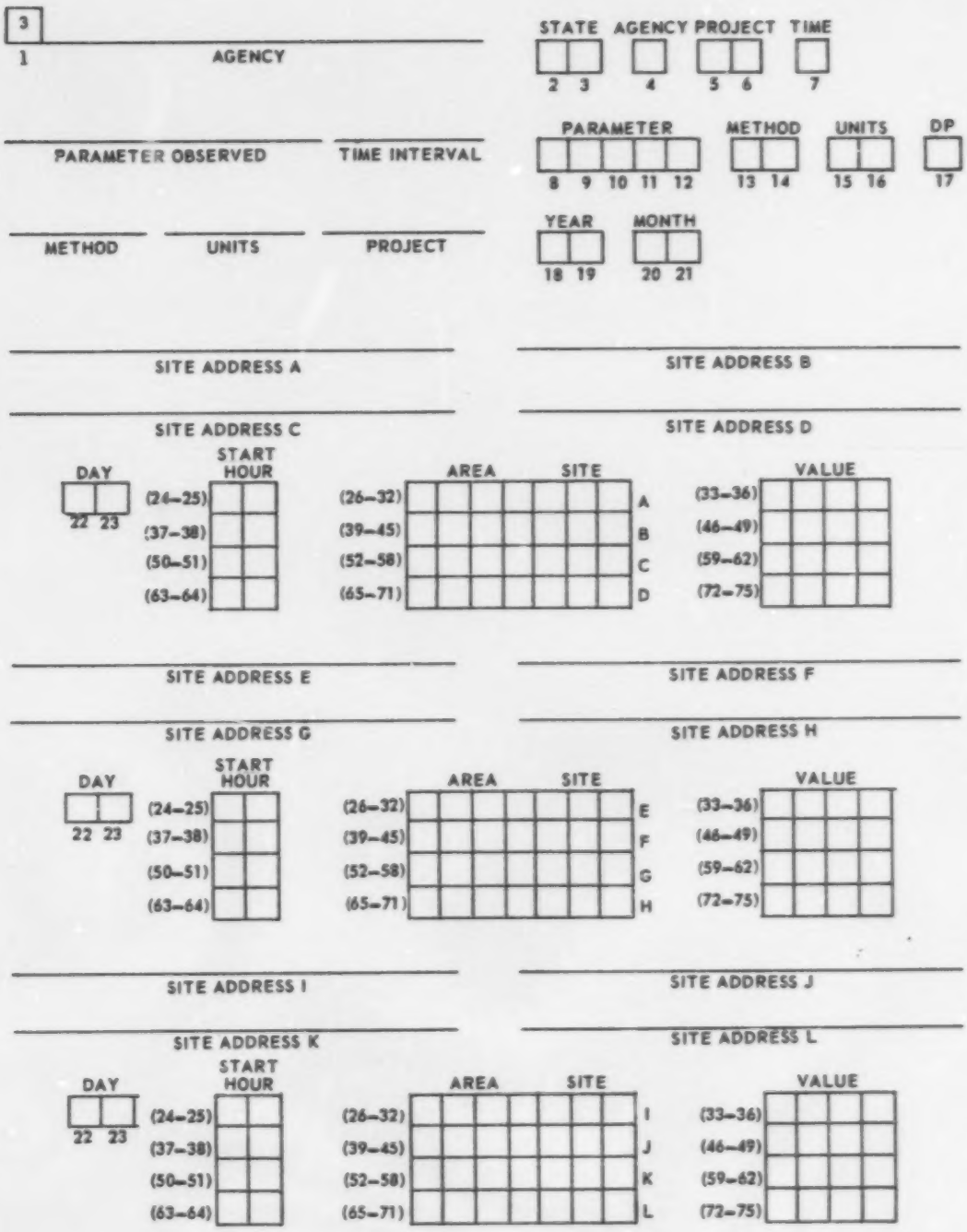


Figure C-2. SAROAD Multiple-Station Data Form.

site codes assigned by EPA for each site address. Note that all sites on one form must be located in a single state. In the last four blocks, record a value. Repeat this process until the 12 lines are filled or all values are recorded. If an agency should determine that none of the forms described in this manual meet specific needs for reporting data, request the help of EPA personnel in developing an appropriate SAROAD-compatible form.

APPENDIX D.
COMPOSITE DATA FORM

Composite data are those values resulting from analyses of a solution containing a number of individual samples that have been chemically combined or "composited." In analyzing for the hi-vol particulate fractions, only a portion of the filter is available. On a small amount of the filter it is possible that not enough of a pollutant is present to be detected. By combining all the filters for a month or quarter, it is possible to determine a value of the pollutant for that time period by analyzing the composite. A significant savings in manpower results from compositing. Since each composite usually combines portions of 5 to 10 individual filters, analysis of the composite provides an approximate value that in effect eliminates the need for 5 to 10 analyses.

The SAROAD Composite Data Form (Figure D-1) is the same as the Daily Data Form for the first 13 columns. A "C" in block 14 (the "time interval") signifies that the card contains composite data. Blocks 15 and 16 contain the last two digits of the year in which the samples were taken. Blocks 17-22 vary by type of composite.

QUARTERLY COMPOSITE

1. In blocks 17-18 code 01-04 to designate the first (January-March), second (April-June), third (July-September), or fourth (October-December) quarter.
2. In blocks 19-20 code the number of individual samples that were composited.
3. Code a 1 in block 21.
4. In block 22 put the interval code of the individual samples that make up the composite (from Code Table 3).

SEASONAL COMPOSITE

1. In blocks 17-18 code 01-04 to designate the first (winter), second (spring), third (summer), or fourth (fall) season.
2. In blocks 19-20, code the nurnber of individual samples that were composited.
3. Code a 2 in block 21.
4. In block 22 put the interval code of the individual samples that make up the composite (from Code Table 3).

MONTHLY COMPOSITE

1. In blocks 17-18 code the month (01-12) in which the composited samples were taken.
2. In blocks 19-20 code the number of individual samples that make up the

ENVIRONMENTAL PROTECTION AGENCY
National Aerometric Data Bank
P.O. Box 12055
Research Triangle Park, N.C. 27711

[2]
1 AGENCY

CITY NAME

SITE ADDRESS

PROJECT

TIME INTERVAL

STATE (2 3) AREA (4 5 6) SITE (7 8 9 10) AGENCY (11) PROJECT (12 13) TIME [C] (14)

YEAR (15 16) PERIOD (17 18) NUMBER OF SAMPLES (19 20) COMP. SAMPLE TYPE TIME (21 22)

FILTER NUMBER	SAMPLE DATE	TOTAL AIR VOLUME M^3	PARTICULATE CONC. $ug\ M^3$

composite.

3. Code a 3 in block 21.
4. In block 22 put the interval code of the individual samples that make up the composite (from Code Table 3).

WEEKLY COMPOSITE

1. In blocks 17-18 code the week (01-52) in which the composited samples were taken.
2. In blocks 19-20 code the number of individual samples that make up the composite.
3. Code a 4 in block 21.
4. In block 22 put the interval code of the individual samples that make up the composite (see Code Table 3).

ANNUAL COMPOSITE

1. In blocks 17-18 code 00.
2. In blocks 19-20 code the number of individual values that make up the composite.
3. Code a 5 in block 21.
4. In block 22 put the interval code of the individual samples that make up the composite (from Code Table 3).

Fill in the remaining part of the card as on the Daily Data Form with one to four different parameter values.

Along the lower left side of the form is space to code values for 20 additional pollutants.

Along the lower right side of the form is space to record pertinent data from the individual samples.

APPENDIX E. EXAMPLES OF ASSIGNING START-HOUR AND LOCATING FIRST OBSERVATION

An agency has the option of choosing any hour within a day to start collecting data. This hour, however, may not be identical to the starting hour that should be recorded on the SAROAD form or the punched cards. The recorded starting hour is the time closest to midnight (00), but still the same day, at which the sampling could have begun and still have given an observation for the actual starting time. For instance, if a 3-hour sampling interval is used starting at 8 a.m., the recorded starting time would be 2 a.m. (02). In recording the data values, the first two 4-digit data values would be skipped and the first recorded observation would begin in boxes (card columns) 43-46.

A convenient rule for calculating the starting hour to be recorded is to subtract the sampling time interval from the actual starting time until a value listed in the appropriate column in Code Table 6 has been reached. This value in Code Table 6 is the appropriate starting value for that time interval. The number of times the time interval was subtracted equals the number of 4-digit values left blank on the SAROAD form (or punched card) before the first data point is entered. To help with this rather confusing item, two examples, along with appropriate illustrations of the completed SAROAD form and punched data card, are presented below. For the illustrations, assume the sampling is on January 1 and 2, 1970, from the population-oriented station at site 012 in city 0960 in state 01 operated by the state agency.

EXAMPLE 1.

Consider an agency that begins sampling at 7 a.m. Sampling is done at a 2-hour interval on an AISI spot tape sampler yielding Cohs/1000 linear feet. To calculate the starting hour (ST HR) that must be recorded to satisfy the SAROAD format, first consult Code Table 6 in the two-hour column. Now, subtract 2 (the sampling interval) from 7 (a.m.) until one of the appropriate starting hours is reached. Only two possibilities are listed under 2-hour sampling: 00 and 01. Two is subtracted from 7 three times before 01 is reached. This now indicates two things: the recorded "ST HR" will be 01 (1 a.m.) and there will be three blank values before the first observation is entered. For the sake of the example, assume the values collected for a particular day were:

7 a.m. - 9 a.m.	= 4.1	5 p.m. - 7 p.m.	= 4.1
9 a.m. - 11 a.m.	= 3.6	7 p.m. - 9 p.m.	= 3.3
11 a.m. - 1 p.m.	= 4.2	9 p.m. - 11 p.m.	= 4.0
1 p.m. - 3 p.m.	= 4.4	11 p.m. - 1 a.m.	= 3.6
3 p.m. - 5 p.m.	= 4.6		

Now refer to Figure E-1a and E-1b to see what the completed SAROAD form and punched card would look like for this day's data.

PRECEDING PAGE BLANK

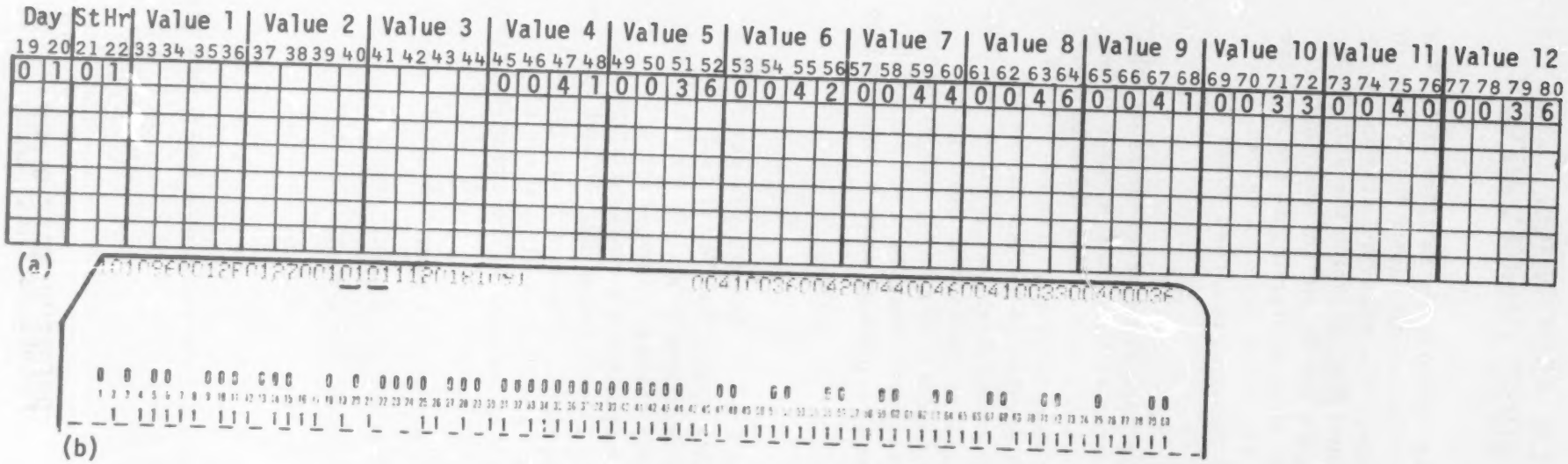


Figure E-1. Example of data coding.

Note that in the space for decimal locator a 1 is entered to indicate one digit to the right of the decimal point.

EXAMPLE 2.

Consider an agency that for some reason does not start sampling until 10 a.m. on a particular day. Assume the pollutant code is 44901, the method code is 96 and the units are 07. Sampling is on a 6-hour interval. Therefore, subtract 6 from 10. This subtraction need be performed only once before an appropriate value, 04, is located in the six-hour column on Code Table 6. The recorded starting hour for the day is 04 and the number of blank values equals 1. For the sake of this example, assume the following data values:

10 a.m.	- 4 p.m.	=	1.04
4 p.m.	- 10 p.m.	=	0.96
10 p.m.	- 4 a.m.	=	0.99
4 a.m.	- 10 a.m.	=	1.12

Note that, although the 10 p.m. - 4 a.m. value runs into the next day, it will be listed with the same day. The next value (4 a.m. - 10 a.m.) will be recorded as the next day's first observation. The recorded starting hour for this next day will be 04, and the value will be located in the first value location. The decimal point for these data will be 2, and it would appear on the SAROAD form and punch card as indicated in Figure E-1c and E-1d.

APPENDIX F.
DETAILED INFORMATION FOR SUBMITTING DATA IN SAROAD FORMAT ON MAGNETIC TAPE

IBM-compatible tapes, which are required for the data bank, are 0.5-inch-wide magnetic tapes containing BCD coded characters recorded at a density of 556 bits per inch or EBCDIC coded characters recorded at a density of 800 or 1600 bits per inch. Internal machine representation of numbers or characters is not acceptable. Floating point, packed decimal, or zoned decimal representation of numerical information should not be written. Agencies not capable of writing this type of tape must submit their data on punched cards or on paper forms.

Agencies capable of providing the type of tape described are encouraged to submit SAROAD-formatted card image records. Use of this standard format eliminates the requirement for special-purpose programming. Agencies using the SAROAD system of coding information who wish to submit data in formats other than card image should contact EPA before sending such tapes so that details can be arranged.

The tape reels should be accompanied by instructions to identify contents of the files, coding and format used, and physical characteristics of the tape (such as density, number of tracks, and parity). Characteristics of individual data files, including sort sequence of data records, should be specified.

CHARACTERISTICS OF DATA FILES

Information on tape is arranged in blocks consisting of one or more records. For example, if card-image formats are used, each record is a single card image and a block might consist of ten card-image records grouped together. Blocking records permits more efficient use of tape. A block cannot be longer than 32,000 characters.

Tapes may contain single or multiple files of information. No label records are written on tapes. Unlabelled tapes contain only data records and tapemarks. Although tapes should be unlabelled, the sender should submit identification with the reel to indicate the tape contents. Individual files on the same tape reel are separated by a single tapemark so that a tapemark follows every data file. An additional tapemark is written so that two tapemarks follow a data file if it is the last or only file on the reel. Extremely large data files could require multiple reels. Following the portion written on all but the last reel a single tapemark is written. On the last reel tapemarks are written just as if only one reel had been required. The organization of data files on one or more reels is shown in Figure F-1.

Table F-1 is a checklist useful for recording the minimum information required to describe a tape. Please include all additional information helpful to persons who are not familiar with the data on the tapes.

PRECEDING PAGE BLANK

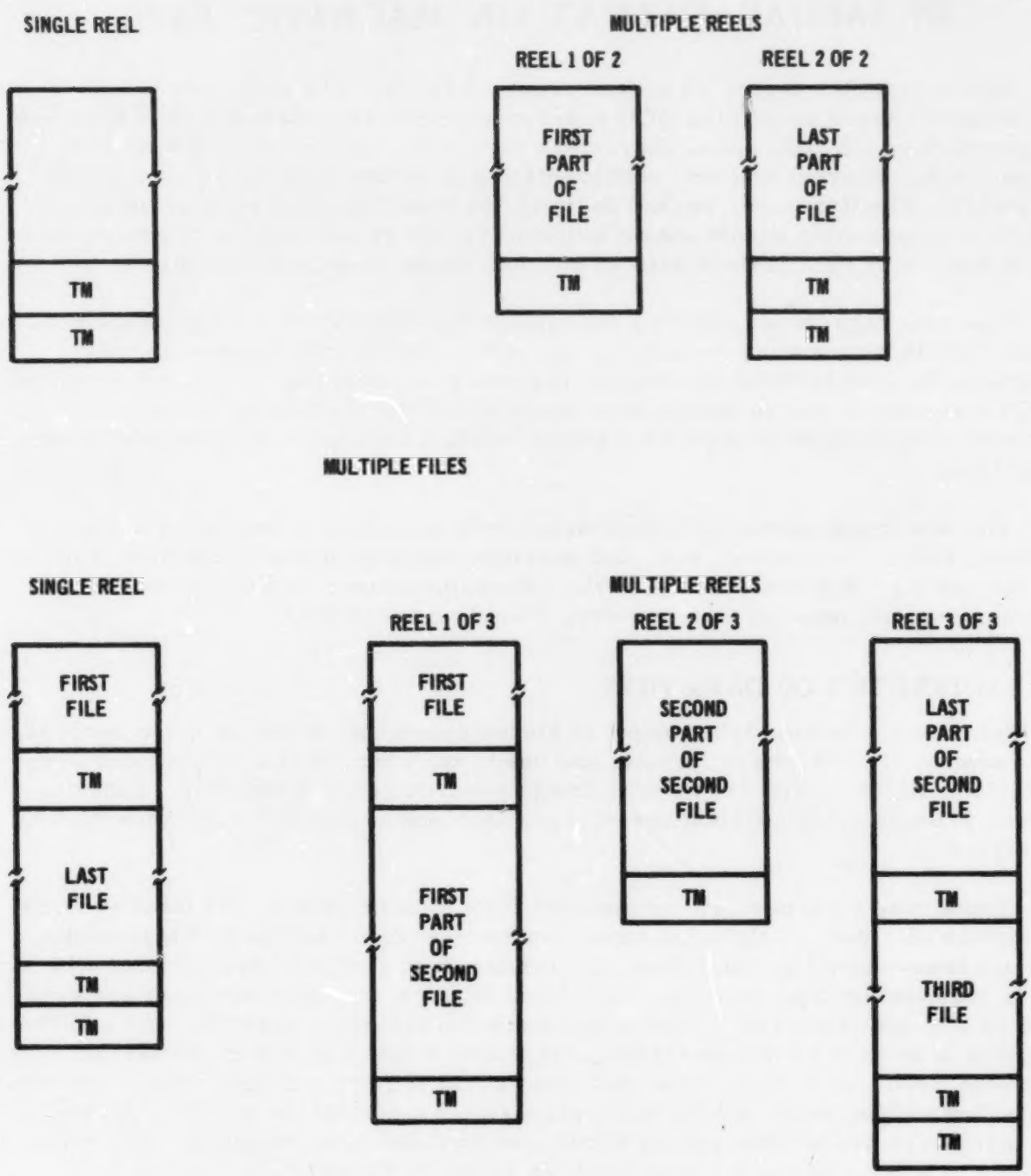


Figure F-1. Organization of files on tapes.

Table F-1. MINIMUM INFORMATION REQUIRED ABOUT TAPE/FILE CHARACTERISTICS

1.	Reel identification	
2.	File number	
3.	Recording density (556, 800, or 1600)	
4.	Number of tracks (7 or 9)	
5.	Record length (in number of characters)	
6.	Blocking factor (in number of records/block)	
7.	Description of file contents	
	- Indicate parameters observed - Number of sites reported - Years spanned - Approximate number of records	
8.	Description of file format	
	- Indicate format identification numbers used and observation time intervals	
9.	Description of file sort sequence	
	- Identify individual fields - Indicate major to minor ordering - For each field indicate if ascending or descending	
10.	Parity (even or odd)	

APPENDIX G.
SUBMITTING PAST DATA IN NONSAROAD FORMAT (FORMS, CARDS, OR TAPE)

It is intended that the National Aerometric Data Bank incorporate the maximum possible amount of air quality data so as to provide an accurate historical record of air quality in the United States. We therefore encourage agencies that cannot provide resources for converting their past records of air quality to SAROAD format to send photocopies of these records to the Data Bank for processing as time allows. When these data are converted to SAROAD formats and entered into the data bank, EPA will provide all converted data to the submitting agency.

Agencies wishing to receive this service are asked to provide certain minimal information along with their records of past data. Of first importance is a completed SAROAD Site Identification form; the significance of accurate information about sampling sites cannot be too much emphasized.

As with current data, an agency may submit past data to the bank in any of three ways: on paper data forms, on punched cards, or on magnetic tape. In whichever form the data are submitted, the agency should provide all the information needed.

Following is a checklist of the specific information required (note that most past records will include many of these items):

1. Completed SAROAD Site Identification form.
2. Pollutants measured.
3. Sampling methods (e.g., Saltzman, West-Gaeke).
4. Instruments.
5. Units in which the measurements are expressed.
6. Decimal point locations (if the decimal point does not explicitly appear in the data tabulations, state exactly how the location of the decimal point can be determined).
7. Date and day of the week on which reading was taken.
8. Time at which reading was taken (specify whether midnight is denoted by 0000 or 2400 hours).
9. Sampling frequency. (How often are readings taken?)
10. Averaging time (duration of sample).
11. Identification of any special codes.
12. Special instructions to persons using data.

PRECEDING PAGE BLANK

PUNCHED CARDS

Agencies submitting data on punched cards should also include:

13. Identification of any unusual formats, characters, or punching.
14. Sequence of data cards (e.g., sort sequence, ordering, arrangement).
15. Identification of machine on which cards were generated.

MAGNETIC TAPE

Note again that the National Aerometric Data Bank can process only IBM-compatible magnetic tape (if the agency cannot submit data on IBM-compatible tape, they should use punched cards or paper forms). Agencies that do submit their past records on magnetic tape should incorporate the following in addition to information on the foregoing checklist of 12 items:

1. Tape characteristics
   a. Tape density in BPI (bits per inch).
   b. Number of tracks (7-track or 9-track, 9-track is preferred).
   c. Type of parity (even or odd).
2. Detailed statement of how tapes are created. (Where applicable, identify utility programs used, include copies of Job Control Language (JCL) listing, or copies of control cards.)
3. Characteristics of data sets
   a. Sequence of the data (i.e., sort sequence).
   b. Length in characters (or bytes) of an individual record.
   c. Number of individual records grouped between Inter Block Gaps (IBGs).
   d. Indication of fixed length or variable records.
   e. Presence or absence of leading or trailing tape, labels, or words (if possible, omit these).
   f. Number of data sets (or files) on the tape.
   g. Approximate number of records in each data set. (Specify whether number is for logical records or physical records.)